Enhancing Investment Management

A Two-Part Project for Enhanced Data Visualization and Intelligent Deal Appraisal

Project Team:

Dante Amicarella: deamicarella@wpi.edu

Sarah LaRusso: sjlarusso@wpi.edu

Maya Liao: mcliao@wpi.edu

Nathan Shemesh: nshemesh@wpi.edu

Project Advisors

Wilson Wong, *Computer Science Department*

Robert Sarnie, *Business School*

Marcel Blais, *Mathematical Sciences Department*

Abstract

In the field of alternative investment management, big data presents challenges and opportunities, necessitating a shift in how firms approach information management. Our team collaborated with an alternative investment firm to reduce the amount of time analysts review potential and historical investment deal data. To achieve this goal, we executed a two-part project: the first part involved constructing a Power BI dashboard of key metrics for their Residential Mortgage-Backed Securities desk using Azure Data Lake, Python, Power BI, and PostgreSQL; the second part utilized the company's proprietary Chat-GPT system and prompt engineering to develop a deal appraisal model using Python, Databricks, Azure Cognitive Search/AI Document Intelligence, and Amazon Textract.

Executive Summary

Investment management, particularly in real estate, plays a pivotal role in shaping financial futures by strategically allocating assets to maximize returns and manage risk for individuals, institutions, or funds. Our team collaborated with an alternative investment company to satisfy two needs: gaining a deeper understanding of key predictors for future mortgage-backed security deals and expediting the review process for corporate consolidation investments. In Power BI, we developed a comprehensive dashboard analyzing historic deals to determine how the value of the deal changed over time. We utilized past deal data as well as external data and proprietary data stored in Postgres and Azure Data Lake. Following the completion of our first project and the acquisition of the alternative investment company by its new parent company, we embarked on a second project centered around the parent company's deal documents on potential company share investments and acquisitions. The goal of this project was to make use of the company's internal GPT-4 large language model and prompt engineering techniques to predict optimal company investments opportunities. This initiative was driven by the need to filter potential investment documents to a smaller subset that reflects worthwhile acquisitions, which will then be sent to the company's Investment Committee for further review. Our collaboration with the company highlights our ability to adapt and efficiently leverage advanced technologies to provide valuable insights and streamlined processes for future investments.

The alternative investment firm manages $73 billion in credit and real estate assets. The company identifies investment opportunities worldwide in global market credit and real estate markets including mortgage-backed securities. Mortgage-backed securities are financial

instruments like bonds, created by consolidating home loans and real estate debt. Banks package and sell these loans to investors, like the alternative investment firm, at a discount. These investors hope to gain value from the loans while mitigating the risk of people defaulting on their mortgages.

In the fall of 2023, the parent company, a global asset manager, acquired the alternative investment firm, expanding the parent company's involvement in credit and real estate, and offering a wider array of investment strategies. The parent company has an internal Investment Committee (IC) in charge of determining which companies to invest in from a sizeable stack of options. Filtering the stack to include only optimal deals would help speed up reading the documents, thus cutting down hours spent on this task. New tools like GPT-4 can be used to quickly filter these deals as the large language model can quickly make sense of large documents. GPT-4 is fine-tuned through Reinforcement Learning from Human Feedback and stands out as an industry leader in language modeling and multimodal AI applications (OpenAI, 2023).

For our first project, regarding mortgage-backed securities, our goal was to create a dashboard that filters the data by each deal, or property. The subsequent steps were taken by our team to achieve our objective:

- *Conducted a thorough analysis of the company's existing data infrastructure, identifying key data sources.*
- *Created new tables in Azure Databricks, connecting public housing data to internal deal data.*
- *Connected the existing and new tables in Power BI.*
- *Built out table views and graphs that drew insights from the data and included a filter to present the data in an accessible and actionable format.*
- *Employed agile practices to ensure fast, iterative feedback and iterations.*

We concluded the project with four completed Power BI dashboards corresponding to five data sources: Redfin, Zillow, ESRI Demographics, New Home Sources, and Multiple Listings Enterprise Solutions. In total there are fifty-one desired graphs or tables. These charts represent a variety of information, such as demographics and locational data, surrounding a mortgage-backed security deal. Example locational graphs include the average selling price of homes in the county of interest and the number of new homes sold year over year. Some demographic visuals showed age group breakdown and household income. Additionally, creating new tables and putting them in Databricks opens the ability for our dashboards in the future to be easily updated and kept current using an existing scheduler in Databricks.

For our second project, we set out with the goal of streamlining the Investment Committee's process of taking potential deals by filtering out 'bad' deals using the alternative investment company's existing GPT-4 model.

- *Created a temporary instance of the company's GPT-4 model, Samantha, using Docker to hold the container and run from VS Code.*
- *Developed initial prompts for the LLM to return key financial metrics from the deal documents.*
- *Collaborated with our sponsor to incorporate a text parser to read the documents as well as implement an image parser to make sense of where words are located on a page.*
- *Determined key financial metrics through document reading and meetings with WPI professors with knowledge of finances.*
- *Put these metrics into a JSON object utilizing Samantha and then pulled values to determine a score for each metric by evaluating buckets that each value could fall into.*
- *Output the score in the user interface and added the pulled values and scores to a table in Postgres.*
- *Analyzed multiple runs of each document to ensure consistency.*

In the end, we developed a scalable and interpretable model, ensuring transparency in decision-making. The basis of the model is an average of sub scores from financial metrics such as net retention, annual recurring revenue growth, and compound annual growth rate. Samantha returns a score on how investable each investment is between 1 and 10 along with the mathematical motivation. To check the reliability of the model, we ran each document through it ten times. The largest range between scores for a single document was 0.86 points and the overall scores returned from all the documents ranged from 7.43 to 10. The backend provides documentation to allow the company to change the bounds for each metric depending on their desired needs. Additionally, the project creates a foundation for how a similar process and prompt can be followed to design a scoring model for other types of documents that the parent company has.

Our project successfully addressed the intricate challenges facing the alternative investment firm in mortgage-backed investments and company investment. We developed two products utilizing a variety of software platforms like Power BI, GPT-4, Docker, DBeaver, and Databricks, while also experiencing what it is like to be in a corporate environment undergoing an acquisition. The developed dashboard and created prompt have empowered the company with the tools needed to make faster, informed, data-driven investment decisions.

Acknowledgements

We would like to express our sincere gratitude to everyone involved in the completion of this project for their invaluable support and contributions. First, we would like to thank our project advisors: Professor Wilson Wong, Professor Robert Sarnie, Professor Marcel Blais, as well as our co-advisor Manasi Danke. We truly appreciate all their guidance, help with overcoming barriers, and direction. We would also like to thank our scrum coach, Marc Trudeau, for his aid with our Jira board and influential enthusiasm for Scrum.

We would also like to thank Professor Kwamie Dunbar. His previous work experience as well as finance knowledge was incredibly valuable to us as he helped us determine which metrics were essential to include in our model given the type of investments we were looking at. He also taught us what these different financial metrics indicated, drastically increasing our finance knowledge, and aiding with the development of our project.

In addition to our WPI team, we would like to extend a huge thank you to our two contacts from the sponsoring company. Our primary contact not only met with us on a daily cadence, but he also rapidly responded to all our Microsoft Team's messages and went above and beyond to complete extra work that helped us accelerate and expand our project. In our daily meetings, he clarified project goals and tasks, provided feedback on what we accomplished, and helped us surpass blockers. The project would not have been the same if not for his continuous devotion!

Table of Contents

List of Figures

List of Tables

List of Equations

Authorship

Section	Main Author(s)	Main Editor(s)
Cover Page	Sarah LaRusso Maya Liao	Dante Amicarella
Abstract	Maya Liao	Dante Amicarella
Executive Summary	Sarah LaRusso	Nathan Shemesh
Acknowledgements	Sarah LaRusso	Dante Amicarella
1.0 Introduction	Maya Liao	Dante Amicarella
2.0 Research	Maya Liao Dante Amicarella Nathan Shemesh Sarah LaRusso	Maya Liao
3.0 Methodology	Dante Amicarella Sarah LaRusso	Nathan Shemesh
4.0 Software Development Environment	Sarah LaRusso Maya Liao	Maya Liao
5.0 Software Requirements	Dante Amicarella Nathan Shemesh	Sarah LaRusso
6.0 Design	Dante Amicarella	Maya Liao
7.0 Implementation	Dante Amicarella Sarah LaRusso Nathan Shemesh Maya Liao	Sarah LaRusso
8.0 Results	Nathan Shemesh Sarah LaRusso	Sarah LaRusso Maya Liao
9.0 Assessment	Maya Liao Dante Amicarella	Dante Amicarella

	Nathan Shemesh Sarah LaRusso	
10.0 Business and Risk Management	Dante Amicarella	Maya Liao
11.0 Future Work	Dante Amicarella Sarah LaRusso	Maya Liao
12.0 Conclusion	Dante Amicarella	Sarah LaRusso

1. Introduction

In the realm of modern finance, investment management emerges as the guiding force that shapes the financial futures of individuals and institutions alike. The overarching objective of investment management is rooted in strategic asset allocation, aiming to maximize returns while effectively managing risk. Of particular interest within this domain is the specialized field of real estate investment management. This sector concentrates on the acquisition, ownership, operation, and disposal of real estate assets. Additionally, it encompasses the professional management of real estate properties and portfolios on behalf of investors, including individuals, institutions, or funds (Simpson, 2022).

Our project involved collaborating with an alternative investment firm to reduce the amount of time analysts review potential and historical investment deal data. Divided into two distinct sub-projects, the first part involved researching the company's historical deal demographic data to create comprehensive dashboards highlighting key performance metrics between deals. Alternatively, the second part of the project focused on utilizing the company's private large language model in the ChatGPT interface and prompt engineering to develop a deal ranking heuristic for appraising existing deal profile documents. Through the remainder of this report, the first project will be referred to as the RMBS Dashboard, and the second project will be referred to as the LLM Deal Appraisal Prompt. The RMBS Dashboard involved use of Python, Microsoft Data Lake, Power BI, and PostgreSQL; the LLM Deal Appraisal Prompt utilized Python, Databricks, Chat-GPT, Azure Cognitive Search/AI Document Intelligence, and Amazon Textract.

2. Research

2.1 Company Background

2.1.1 The Alternative Investment Company

Established in 1988, this New York City-based alternative investment management firm boasts a workforce of over 650 employees across offices in the United States, Europe, and Asia. Managing approximately $73 billion across diverse credit and real estate assets, the company has consistently adhered to a disciplined approach to portfolio construction, placing a strong emphasis on capital preservation throughout its history. By conducting extensive research into industry and company fundamentals, the firm identifies global investment opportunities and exploits inefficiencies in credit and real estate markets worldwide. The company is structured around two investment strategies: credit and real estate. These strategies guide the design of portfolios aimed at maximizing returns on investments (ROI) while prioritizing risk management.

The company's real estate sector comprises three subdivisions: Global Private Equity Real Estate, Commercial Real Estate Debt, and Net Lease Real Estate. In contrast, its credit branch includes Corporate Credit, Structured Credit, Middle Market Direct Lending, and Multi-Strategy. The real estate division primarily focuses on commercial mortgage-backed securities in the United States. Residential Mortgage-Backed Securities (RMBS) involve consolidating residential mortgages to create securities backed by mortgage payments.

Alternative investment management firms employ various strategies when dealing with RMBS, depending on their investment goals and risk tolerance. The company's RMBS team aims to identify pricing inefficiencies by conducting fundamental analyses of loans associated with properties and related risks. Similar firms utilize strategies such as buy and hold, credit

analysis, yield enhancement, active trading, securitization and structuring, hedging, relative value, non-agency RMBS, macro factors, and quantitative models.

It is crucial to note that the chosen RMBS strategy significantly affects the risk and return profile of an investment portfolio. As a result, these strategies can change with industry regulations and market dynamics. Consequently, alternative investment firms must carefully assess market conditions and investor objectives to determine the most suitable approach for their clients or funds.

2.1.2 The Parent Company

Founded in 1992, the alternative investment management firm's parent company is a prominent global alternative asset manager currently overseeing a substantial $212 billion in assets under management. With a robust presence spanning 30 countries, the parent company effectively manages a portfolio of over 300 active companies worldwide. Operating similarly to many investment firms, the parent company deploys its investments across an extensive array of strategies, encompassing private equity, impact, credit, real estate, and market solutions. Its distinctive corporate approach centers around sector efficiency. Within the parent company's framework, emphasis is placed on investors collaborating across diverse products and strategies. This collaboration aims to cultivate profound insights and expansive networks within their respective key sectors. These sectors actively engage in the exchange of ideas and intellectual capital, fostering the creation of comprehensive cross-platform capital solutions.

In Fall 2023, the parent company made a significant announcement regarding its successful acquisition of the alternative investment company. Under the new arrangement, the alternative investment company will operate as a subsidiary of the parent company, functioning as a diversified credit and real estate investing platform valued at $74 billion. This strategic

acquisition serves as a pivotal step towards expanding the parent company's involvement in credit expansion and real estate, thereby offering a wider array of alternative investment strategies to clients. This partnership not only signifies a substantial growth opportunity but also fosters the integration of over 350 compatible institutional limited partnerships. The union between these entities is expected to unlock substantial prospects for revenue growth while enabling the optimization and scaling of operations.

2.2 Mortgage-Backed Securities (MBS)

Mortgage-backed securities (MBS) are investment instruments akin to bonds that involve the consolidation of a collection of home loans and other types of real estate debt. Banking institutions are responsible for packaging these loans to sell at a discount to investors as a type of collateralized bond (Kenton, 2023). This allows investors to make a profit as they will acquire the interest that homeowners pay towards their mortgages. MBS essentially turns banks into intermediaries between homebuyers and investors. The banks do this so they can make a profit without having to risk homeowners possibly defaulting on their payments sometime in the future (Kagan, 2023).

MBS often involves the use of tranches, which are segmented securities—usually bonds or mortgages—categorized by risk, maturity, or other factors to attract diverse investors with varying preferences (Chen, 2020). Tranches are prevalent in securitized debt products like collateralized debt obligations (CDOs) and collateralized mortgage obligations (CMOs). CDOs are instrumental in pooling together a collection of cash flow-generating assets. As shown in Figure 2.2, an MBS is constructed from multiple mortgage pools, each containing a wide variety of loans, ranging from safe loans with lower interest rates to riskier loans with higher rates. These loans can be categorized based on their expected return, with riskier unsecured loans

yielding higher returns and safer senior secured loans resulting in lower returns. These distinct mortgage pools each have their unique maturity timelines, which play a significant role in determining the associated risk and reward aspects. Consequently, tranches are created to segment these various mortgage profiles into slices with financial terms tailored to the preferences of specific investors. CMOs are structured with several tranches organized by their risk profiles (Kagan, 2023).

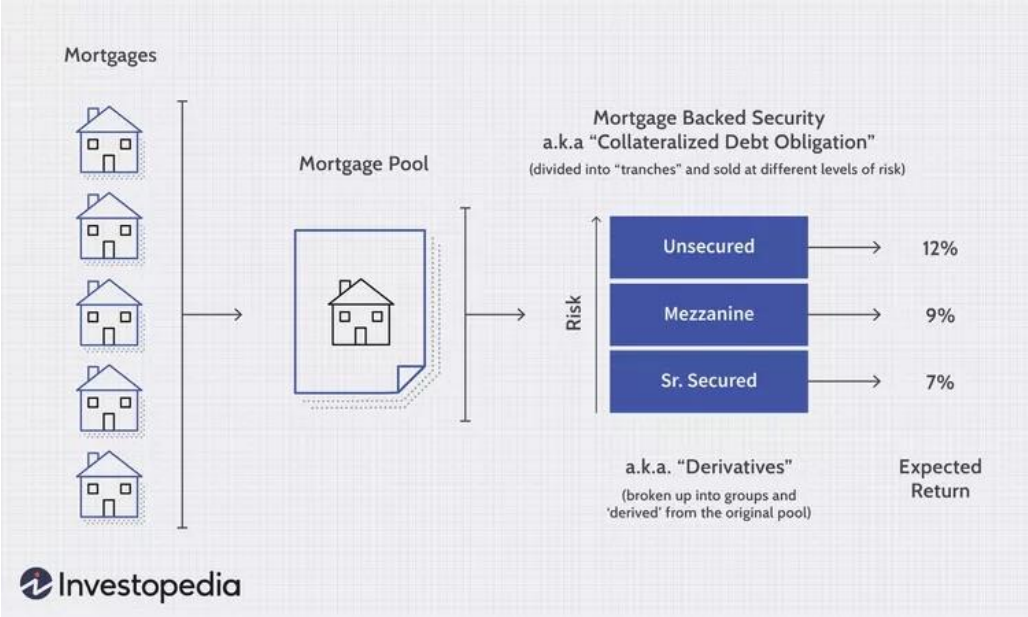


Figure 2.2: Mortgage-Backed Securities and Tranches (Jiang, 2020)

2.3 Large Language Models

2.3.1 Overview

Large language models (LLMs) are a category of deep learning structures designed to "predict the probability of sentences, paragraphs, or even entire documents" (Google, 2023). These models are trained on extensive amounts of data and utilize a set of neural networks called a transformer. The transformer consists of encoders and decoders that evaluate context and

semantics by tracing connections within sequential data, similar to how they identify the relationship between words in a sentence (Nvidia, 2023).

Traditionally, artificial intelligence models were primarily focused on perceiving and understanding information. However, the advent of large language models (LLMs), trained on extensive internet-scale datasets containing hundreds of billions of parameters, has revolutionized AI's capability to create human-like content. These models exhibit skills in reading, writing, coding, drawing, and creative content generation, significantly enhancing human creativity and productivity in various industries (Nvidia, 2023).

2.3.2 Chat-GPT

ChatGPT is a version of the Generative Pre-trained Transformer (GPT) series developed by OpenAI, specifically designed for natural language processing and generation in conversational contexts. It is an AI language model that uses deep learning to comprehend and generate human-like text in response to a variety of user prompts. ChatGPT is trained on a vast amount of information from the internet, allowing it to "answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests" (Introducing ChatGPT, 2022).

Since its release on November 30, 2022, ChatGPT has been adopted across a wide variety of industries, including but not limited to healthcare, finance, and entertainment. For example, Duolingo, a company specializing in language education, has introduced two novel functionalities powered by GPT-4, the most advanced among OpenAI's ChatGPT models (Marr, 2023). With Duolingo Max, language learners receive detailed explanations in conversational language, akin to guidance from a human tutor, which can detail why their responses to practice or test questions were right or wrong (Marr, 2023). The second feature offers students an

opportunity to enhance their language skills through role-playing conversational scenarios with AI-generated characters (Marr, 2023).

2.3.3 The Math Behind LLMs

LLMs are built on neural network transformer models to utilize very large datasets to perform natural language processing tasks like recognizing and generating text. LLMs utilize multiple transformers to perform different layer tasks as they get a deeper understanding of the input data (Lee, 2023). Transformers first take in the input and break it into units called tokens, then they mathematically discover relationships between the tokens. Additionally, these models have self-attention, which weighs importance to different parts of the input to understand relationships between parts.

During the first stage, a transformer will take in the input sentence and split it by word. It then determines the list of numbers in a large dimensional space representing this word, also known as a word vector (Lee, 2023). Figure 2.4.1 shows how words can be broken into a word vector and then gives a representation of how that may look in space. In contrast to the figure, these word vectors exist in a space with hundreds of dimensions (Lee, 2023).

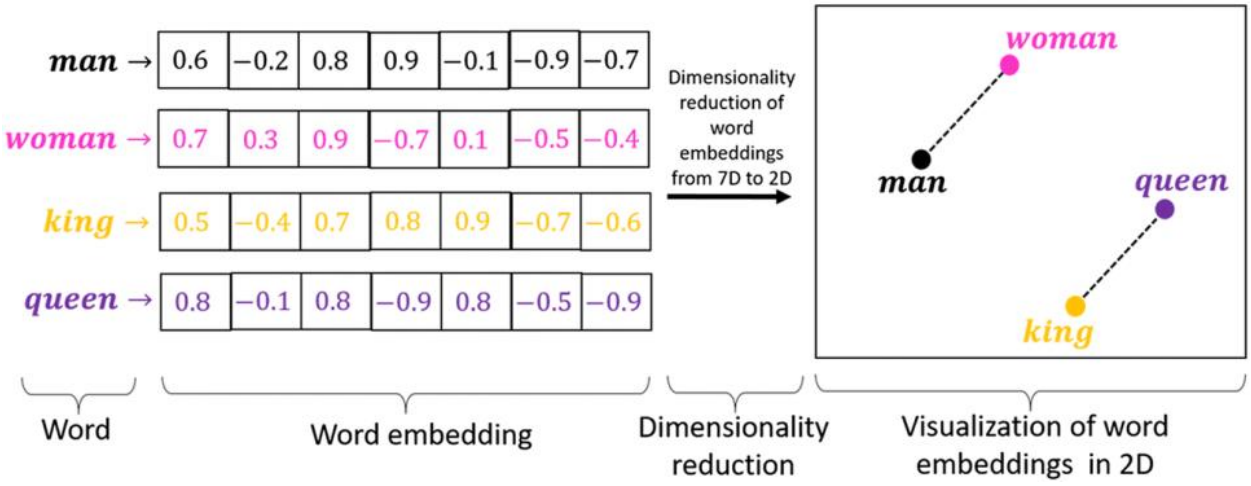


Figure 2.3.1. Representation of an English word embedded as word vector and visualized in space (datascience904, 2019).

The word is represented by a vector of numbers rather than letters as mathematical operations can be performed on vectors to contextualize the word's location relative to other words in space (Lee, 2023). To consider how close two vectors are, one can calculate the Euclidean distance between the two vectors (Equation 1). Additionally, one can calculate the direction and magnitude between two vectors. For example, consider the word vector for 'dog' and subtract that for 'puppy' to discover how closely related the two are to one another and in what direction as displayed by Equation 2. Then this vector can be applied to another word to determine which word has a similar relationship. For instance, one can add this vector to the word 'cat' to find a vector with a similar relationship in space. Then utilizing the result, the model would choose the word closest to this vector in space, which in this case we would expect to be 'kitty'.

$$Euclidean\ Distance = \sqrt{\sum_{i=1}^{n} (\hat{u}_i - \hat{v}_i)^2}$$

Equation 1: Euclidean distance between two vectors u and v of an n-dimensional space

$$dog = [0.7, 0.3, -0.5, 0.2]$$
$$puppy = [0.4, -0.2, 0.1, 0.1]$$
$$dog - puppy = [0.3, 0.5, -0.6, 0.1]$$

Equation 2: Example of operations to compare word vectors

After the transformer represents words in space, it uses a mechanism known as self-attention to capture dependencies between words and prioritize certain values (Karami, 2023). Three components needed for self-attention are as follows:

- *Vector Values (V) – the content inputted into the LLM*
- *Vector Query (Q) – one piece of information the model is currently querying*
- *Key Vectors (K) – another input that the query is referencing against*

As an example, consider the sentence "My friend's dog ate his homework." Let the word 'his' be the vector we are currently querying and each of the words in the sentence become key vectors. The query vector will be a representation of 'his' that highlights the characteristic it is looking for, which may be a masculine noun to pair this pronoun with (Lee, 2023). Then when we look through all the key vectors which are representations of the other words in the sentence. For example, the key vector for 'dog,' will provide information stating this may be a masculine noun. We would then find a weight using the query vector and key vector to assign to the 'dog' value vector. An example of comparing the query vector with key vectors is shown in Figure 2.4.2 below.

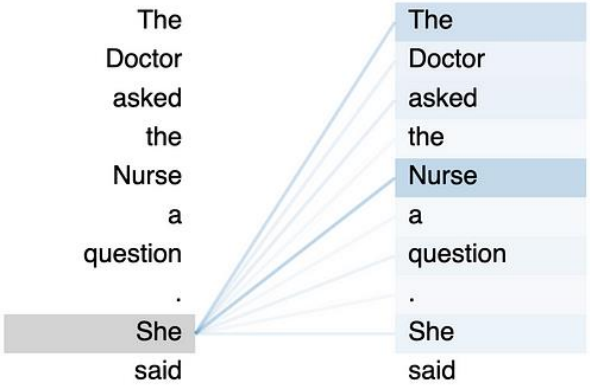


Figure 2.3.2: Comparison of a query vector (left) with key vectors (right) (Morgan, 2023).

To determine the weight for each value vector, transformers use "scaled dot-product attention" (Karami, 2023). First, one must calculate the dot product of each query-key pair representing how well they match one another; we demonstrate this calculation in Equation 3. This will be done for each key while maintaining the same query vector, and results will be

placed into a vector of length equal to the size of the number of keys (Equation 4). Vectors that point in similar directions will result in a larger dot product. Next, the transformer scales the score by dividing it by the square root of the number of elements in the query vector to prevent overgrowth for high dimensions (Karami, 2023). Then, we apply the SoftMax function to the vector to ensure that we have positive weights that sum to one (Equation 5). Finally, we perform a weighted sum of the value vector as shown in Equation 6 to obtain the output for the chosen query vector (Karami, 2023).

$$\boldsymbol{q} \cdot \boldsymbol{k} = \sum_{i=1}^{n} \boldsymbol{q}_i * \boldsymbol{k}_i$$

*Equation 3: Given a query vector **q**, a key vector **k**, and n dimensions, the dot product is the sum of the products of pairwise elements*

$$dots = [\, \boldsymbol{q} \cdot \boldsymbol{K_1}, \boldsymbol{q} \cdot \boldsymbol{K_2}, \boldsymbol{q} \cdot \boldsymbol{K_3}, \ldots, \boldsymbol{q} \cdot \boldsymbol{K_m}]$$

*Equation 4: The first step for attention calculation, given a query vector **q** and m key vectors labeled $\boldsymbol{K_1}$ through $\boldsymbol{K_m}$*

$$weights = SoftMax(\frac{dots}{\sqrt{n}})$$

Equation 5: A weight vector of size m generated after performing SoftMax

$$Attention = \sum_{i=1}^{m} weights_i * V_i$$

Equation 6: The formula for attention, given calculated weights and m value vectors labeled $\boldsymbol{V_1}$ through $\boldsymbol{V_m}$

2.3.4 GPT-4

Launched on March 14, 2023, GPT-4, officially known as Generated Pretrained Transformer 4, represents a pioneering achievement in the domain of AI. This multimodal marvel, developed by OpenAI, transcends conventional language models, showcasing an

unparalleled level of proficiency in various professional and academic fields. Its unique capability to process both text and images sets a new standard in the realm of AI technology, enabling a more holistic understanding of information (*Introducing ChatGPT*, 2022).

GPT-4 is described as, “a Transformer-based model pre-trained to predict the next token in a document... fine-tuned using Reinforcement Learning from Human Feedback” (OpenAI, 2023). Rigorous testing has solidified its superiority over prior models and most cutting-edge systems, establishing GPT-4 as an industry leader in language modeling and multimodal AI applications. Despite its remarkable capabilities, GPT-4, like any AI system, comes with its set of limitations and safety considerations. Understanding and acknowledging these boundaries is crucial when utilizing this extensive language model for knowledge extraction and various applications.

2.4 Prompt Engineering

The enhanced capabilities of LLMs in today's world have given rise to a new discipline known as prompt engineering. In this field, developers optimize prompts for efficient use in LLMs across various applications. Prompt engineering, in essence, involves formatting a question or instruction to yield a specific output that adds value to the user.

When creating a prompt, four elements merit consideration. The first is instructions, specifying the task the model should perform. The second is context, external information guiding the model to a better response. The third is input data, representing the user's question or goal. The final element is the output indicator, describing how the user wants the output organized. While it is not mandatory to include all four elements in an LLM prompt, judicious use of these steps helps avoid impreciseness in the model and promotes a more focused approach, leading to improved outputs (Saravia, 2023).

There are various techniques used when creating prompts, and one prominent method is generated knowledge prompting. The general idea behind this approach can be divided into two separate tasks. The first task is to query the LLM to provide information about the topic in question, and the next task is to use the generated information and concatenate it with a question to produce an answer. Going through these two steps can enhance the precision of the LLM's internal logic. However, there are some potentially fatal issues with the generated knowledge approach, as data can be outdated or incomplete, and the model cannot verify the source of the information (Liu et al., 2022).

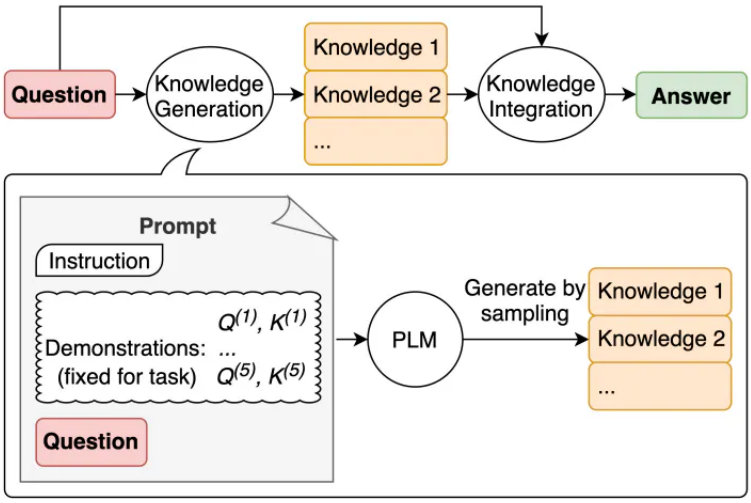


Figure 2.4: General knowledge prompting (Liu et al., 2022).

For prompts focused on exploration, the tree of thought (ToT) approach is effective. ToT maintains a tree structure where each thought serves as an intermediate step toward problem-solving. This allows an LLM to assess progress through deliberate reasoning. By combining the LLM's ability to generate and evaluate thoughts with search algorithms, systematic exploration occurs, including lookahead and backtracking. This enables the LLM to consider multiple potential outputs and select the best one (Saravia, 2023).

General-purpose LLMs can perform various tasks, but they encounter challenges when assigned more complex knowledge-based questions. Fortunately, it is possible to create LLMs with access to an external knowledge base; however, these models require retraining each time the knowledge base is updated with new information. Researchers at Meta have developed a method called Retrieval Augmented Generation (RAG), which combines the advantages of knowledge base models and general-purpose models (Riedel et al., 2020).

RAG consists of two components: informational retrieval and a text generator. When given an input, RAG retrieves the necessary documents from the knowledge base. These documents are then concatenated with the original input prompt and processed through the text generator to produce the final output. The benefits of RAG include the ease with which internal knowledge can be modified or supplemented as needed without requiring the entire model to be retrained (Riedel et al., 2020).

2.5 Investment Deal Performance Metrics

2.5.1 Annual Recurring Revenue

Annual Recurring Revenue (ARR) encompasses the total anticipated revenue generated continuously by a product or business, forecasted for a period of one year (ProductPlan, 2023). This type of metric is typically used by Software as a Service (SaaS) and other subscription-based companies, as it offers a more precise depiction of the company's ongoing revenue compared to Monthly Recurring Revenue (MRR). This is because ARR provides a “longer-term view of your company's revenue. You can use it for reporting, forecasting, valuation, and investor communication...it provides a stable and predictable measure of your company's revenue” (Kipfolio, 2023). In simpler terms, ARR represents the complete revenue a company

anticipates from customers committed to renewing their subscriptions at the existing rate for the subsequent year (Verlaque, 2023).

$$ARR = Monthly\ Recurring\ Revenue \cdot 12$$

Equation 7: Annual Recurring Revenue

2.5.2 Annual Recurring Revenue Year-over-Year Growth

ARR Year-over-Year (YoY) growth is a metric that signifies the variation in annual recurring revenue during a specified timeframe, typically expressed as a percentage. A consistent rise in the ARR growth rate from one year to the next often signals a strong alignment between the product and the market, suggesting a positive product-market fit (*ARR Growth Rate*, 2023). YoY ARR growth is calculated by comparing the ARR values of two consecutive years as follows:

$$\%\ YoY\ ARR\ Growth = \frac{(Current\ Year\ ARR - Previous\ Year\ ARR)}{Previous\ Year\ ARR} \cdot 100$$

Equation 8: Percent Year-over-Year ARR Growth

This calculation expresses the percentage change in ARR from the previous year to the current year, providing insights into the growth or decline of the ARR metric over that period. A positive YoY ARR growth indicates an increase, while a negative value suggests a decrease in ARR from one year to the next (Frankel et al., 2023).

2.5.3 Revenue Retention

Net Revenue Retention (NRR), also known as net dollar retention (NDR), evaluates the variation in revenue generated by a specific group of customers between different time periods. NRR considers the revenue lost due to customer attrition or downgrades, along with any augmentations in total revenue (whether it be annual recurring revenue or monthly recurring revenue) over a defined duration (Groccia, 2022).

$$NRR = \frac{(Starting\ ARR\ +\ Expansion\ in\ ARR\ -\ Contraction\ in\ ARR)}{Starting\ ARR} \cdot 100$$

Equation 9: Percent Net Revenue Retention

In short, NRR helps indicate the extent to which a product is valued by existing customers and their overall satisfaction with business strategies. A high NRR signifies successful customer retention and the ability to generate additional revenue from existing customers, showcasing the potential for sustained growth without heavy reliance on acquiring new customers. Beyond its financial implications, NRR can also be indicative of effective pricing strategies and reflects the company's resilience in withstanding customer attrition and market fluctuations. Investors and stakeholders often use NRR as a key indicator of a subscription-based business's financial stability and long-term sustainability, making it an essential metric for assessing the success and health of such enterprises (Stripe, 2023).

Alternatively, not to be confused with NRR, Gross Revenue Retention (GRR) gauges the percentage of recurring revenue retained from existing customers over a set period, excluding any additional revenue from upsells or expansions. It focuses on potential revenue loss, known as "churn," from customers downgrading or canceling subscriptions (Stripe, 2023).

$$GRR = \frac{(Ending\ ARR + Change\ in\ ARR)}{Starting\ ARR} \cdot 100$$

Equation 10: Percent Gross Revenue Retention

Unlike NRR, GRR focuses “solely on the negative impacts, specifically the revenue lost due to downgrades or customer churn. It does not factor in any upsell or expansion revenue” (Stripe, 2023). This means that GRR provides a more conservative measure to aid businesses in understanding their success in retaining customers at existing subscription levels.

2.5.4 SaaS Magic Number

Magic number is a metric for Software-as-a-Service (SaaS) subscription companies that poses the question, "For each dollar spent acquiring new customers with sales and marketing (S&M) efforts, how many dollars' worth of revenue do we create for the company?" (Mosaic, 2023). To calculate the magic number, quarterly revenue (ARR) and customer acquisition cost (CAC) are required, as shown in Equation 11.

$$SaaS\ Magic\ Number = \frac{(Current\ Quarter\ ARR - Prior\ Quarter\ ARR)}{Previous\ Quarter\ CAC}$$

Equation 11: Magic Number

This metric measures the efficiency of sales and marketing at a company in a single quarter. A magic number of 1 means the company's additional earnings pay off the amount they spent on CAC. A low magic number is an indicator that the subscription cost may be too high or does not align well with the audience, while a magic number greater than 0.75 suggests a good marketing strategy.

Figure 2.5: Magic Number Target Ranges (Mosaic, 2023)

2.5.5 Gross Margin

Gross margin refers to the measure between the gross profits, which are the profit the company makes after subtracting the cost of producing its product, compared to the revenue or sales of a company's product (Bloomenthal, 2022). The gross margin is expressed as a percentage and the higher the margin is the more capital the company is able to retain; however, the gross margin mainly focuses on revenue and the cost of goods sold and does not take into account all of the business's expenses.

$$Gross\ Margin = \frac{(Revenue - COGS)}{Total\ Revenue} \cdot 100$$

Equation 12: Gross Margin

The gross margin allows companies to measure how their production cost relates to their revenues, and they may adjust production by cutting labor costs or finding cheaper materials. They could also increase the cost of the product. In the end, the gross margin is just one metric for an investor to determine whether a company is a good investment.

2.5.6 Contracted Annual Recurring Revenue

Contracted Annual Recurring Revenue (CARR) is a financial metric used in the Software-as-a-Service (SaaS) industry referring to the predictable, recurring stream of revenue generated from subscription-based contracts annually. In simpler terms, it represents the sum of subscriptions from its customers each year. CARR is a valuable metric as it measures "revenue stability, growth potential, and overall performance" of a SaaS company (ClouldBlue, 2023). To calculate CARR, a company must determine all active contracts, annualize monthly contracts by multiplying by 12, and sum all of them together as shown in the formula below.

$$CARR = \sum(annualized\ value\ of\ active\ contracts)$$

Equation 13: CARR

2.5.7 Year-over-year Contracted Annual Recurring Revenue Growth

% YoY CARR growth is a measure of the change in a company's contracted revenue from the previous year. It is calculated as the current year's CARR minus the previous year's, all divided by the previous year's as shown in Equation 14. A positive value denotes an increase in CARR from the previous year and indicates company growth.

$$\%YoY\ CARR\ Growth = \frac{(CARR_{curr} - CARR_{prev})}{CARR_{prev}}$$

Equation 14: %YoY CARR Growth

2.5.8 Revenue

Revenue is the total income generated from the sales of goods or services related to a company's operations. Revenue, which can also be known as gross sales, sits at the top of any company's financial statement (Boyte-White, 2023). This "top line" value is calculated prior to

any expenses being accounted for demonstrating a business's ability to sell their goods or services.

$$Total\ Revenue = \ \#\ of\ units\ sold * price\ per\ unit$$

Equation 15: Total Revenue

YoY Revenue growth, measured by subtracting the current year's total revenue from the previous year's total revenue and multiplying it by 100 demonstrates a positive or negative result. Over the course of the long term, it is important for companies to be able to track and compare their revenue growth year to year to see how they are faring against market trends, prices of raw materials, employees, and other fluctuating metrics to see how they can further maximize profits (Ascent, 2022).

2.5.9 Total Addressable Market

When gauging whether to create a product or service, it is important to understand the market you are entering and the size of it. Market size refers to the total number of potential buyers for your product or service, so it an upper limit to a company's total revenue potential (Ghanizada, 2023). This metric determines whether there is enough demand for a company to enter their target market. To understand how to size the market, a company must take into consideration the total addressable market (TAM). TAM refers to the maximum size of opportunity for a particular service or product and can be calculated by using the average revenue per user (ARPU) multiplied by the total potential customers in market (ProductPlan, 2023).

$$TAM = ARPU * total\ potential\ customers\ in\ market$$

Equation 16: TAM

Despite there being one agreed formula, the method to calculate TAM can either be done in a bottom-up or a top-down approach. Each method's goal is to quantify the total potential customers in the market through different scopes.

2.5.10 Compound Annual Growth Rate

The Compound Annual Growth Rate (CAGR) is the rate of return for an investment to grow from start to finish (Fernando, 2023). This measure considers compounding throughout the investment and is often used to measure and compare past performance to projected future returns (CFI, 2023). CAGR should not be mistaken for a true return rate, rather it is a rate at which an investment would have grown if the rate remained consistent every year and profits were continuously reinvested at the end of each year.

$$CAGR = \left(\left(\frac{EV}{BV} \right)^{\frac{1}{n}} - 1 \right) * 100$$

$$n = Number\ of\ years$$
$$EV = Ending\ Value$$
$$BV = Beginning\ Value$$

Equation 17: Percent Compound Annual Growth Rate

Another advantage of CAGR is that it produces a geometric average which can be used to compare different investment types to one another. This cost analysis would allow companies to determine which investments are most profitable without understanding the volatility of the stock.

2.5.11 Adjusted EBITDA

The earnings before interest, taxes, depreciation, and amortization (EBITDA) serves as a tool to estimate pre-tax cash flow from operations meaning that this measure can determine a company's profitability based on the transactions that have taken place between other businesses of similar size that are in the same industry (O'Dore, 2023). For EBITDA comparisons to be accurate it needs to be adjusted to account for any possible anomalies that may distort the EBITDA since different companies may have expenses that are unique to them.

$$\ni + IT + DA = EBITDA$$

$$EBITDA \pm A = AdjustedEBITDA$$

$$where:$$

$$IT = \ Interest \ \& \ taxes$$

$$DA = \ Depreciation \ \& \ anorization$$

$$A = Adjustments$$

$$\ni = \ "contains \ as \ member"$$

Equation 18: EBITDA (Kenton, 2022a)

Some of the most common EBITDA adjustments include unrealized gains or losses, non-cash expenses, litigation expenses, gains or losses on foreign exchange, non-operation income, and share-based compensation. EBITDA is an effective metric for investors to use because it is a normalized value it does not suffer from fluctuations due to different accounting methods or capital structures, and investors can make more straightforward comparisons between companies (Scott, 2022).

2.5.12 Total Enterprise Valuation

Total Enterprise Valuation, or TEV, measures the total value of a company accounting for market capitalization, debt, preferred stock, and cash equivalents (Kenton, 2022b). It is used

to value companies with varying levels of debt. It is commonly used when assessing the health of a target company for a takeover or merger and is a better assessment than market capitalization alone as it accounts for other factors.

$$TEV = \text{Market Capitilization} + \text{Market Value of Debt} + \text{Preferred Stock} - \text{Cash Equivalents}$$

Equation 19: TEV

The equation for TEV is shown above and contains the following parts:

- *Market capitalization - The total market value of a company's outstanding shares.*
- *Total debt – Sum of company's remaining short-term and long-term debt.*
- *Preferred stock – ownership in a company similar to stock but with fixed dividends and higher claims and earnings.*
- *Cash Equivalents – cash or easily convertible instruments such as securities and treasury bills*

When comparing two companies with similar market capitalization, an investor may want to choose the company with a smaller TEV as they would not have to pay as much in debt in addition to its assets.

3. Software Development Environment

Throughout the project, the team employed an array of software packages and tools to facilitate our work. The company had a vast set of pre-existing infrastructure and tools to process and store data. Our team leveraged the company's existing infrastructure and introduced external software as needed to complete the project tasks.

3.1 Data Processing Infrastructure

3.1.1 Microsoft Power BI

Microsoft Power BI is a visualization tool that allows users to build interactive dashboards. Some abilities include connection to data sources, transformation of data, and data visualization creation. The user can select a variety of visualizations from the Visualizations panel including 'line chart,' 'clustered column chart,' and 'table.' There is also a 'slicer' object to filter the data in the charts by a particular column (Microsoft, 2023).

3.1.2 Apache Spark

Apache Spark is an open-source distributed computing system that is designed for big data processing and analytics. It employs in-memory caching and optimized query execution to enable quick analytic queries on data batches of varying sizes. Moreover, by offering development APIs in Java, Scala, Python, and R, it supports code reuse across a spectrum of tasks like real-time analytics, machine learning, and graph processing (AWS, 2023).

3.1.3 Docker

Docker Desktop is a product that provides a system for running applications in an isolated environment called a container. These containers enable users to create a local instance of an application. Subsequently, users can make changes locally, rebuild, and run the container to verify those changes (Docker, 2023). We utilized Docker to run an unpublished version of the company's LLM, so as not to make changes to their working and running version used company wide. We wrote our code in VS Code and used the terminal to build and run the Docker instance.

3.1.4 DBeaver

DBeaver is a SQL database administration tool that provides an interface to navigate through a database system, open different schemas, and view the relations or tables within each. It presents each table in an easy-to-read format. Users can also write SQL code in DBeaver to interact with the added tables (DBeaver, 2023). We primarily used the tool to see the names of each table, where they were stored, and to gain an understanding of the field titles corresponding to each.

3.2 Project Management Software

3.2.1 Trello

The team used Trello for sprint retrospectives. Trello is an organizational tool for making lists that allows users to create a board, add lists or columns to that board, and populate these lists with items (Atlassian, 2023). We established four columns: 'what went well,' 'what did not go well,' 'surprises/challenges,' and 'ideas/next steps.' Each team member contributed to these columns, and we later regrouped to discuss the entries.

3.2.2 Jira

Jira is project tracking software created by Atlassian. We used the software to manage our sprint board and backlog, as well as to keep track of story points accomplished in each sprint. The software can track various issue types, including tasks, stories, subtasks, bugs, and epics (Atlassian, 2023).

3.2.3 Azure DevOps

Azure DevOps is a Microsoft product offering several capabilities including version control (Microsoft Azure, 2023). Our team used the product to share code updates with one another and maintain a history of past versions. We employed Git in the VS Code terminal to push and pull our files to and from the DevOps cloud.

3.3 Programming Environment

3.3.1 Visual Studio Code (VS Code)

The team utilized Visual Studio code, or VS Code, as the integrated development environment (IDE) for our project (Microsoft, 2023b). VS Code has features to view coding files, run your files, debug, and can interact with git. We used it to hold and edit Python files for both projects we did.

3.3.2 Databricks

Databricks is a “unified, open analytics platform for building, deploying, sharing, and maintaining enterprise-grade data, analytics, and AI solutions at scale” (Databricks, 2023). The platform was founded by the creators of Apache Spark, and it aims to simplify the complexities

of big data processing and make it more accessible to a broader audience. We utilized Databricks as a backend Python code editor for data engineering, integration, and exploration tasks in our project.

3.4 Data Sources

In our initial epic, we utilized a comprehensive array of data sources, combining information from Redfin, the Census Bureau, Zillow, and datasets obtained through external channels and web scraping. Within the firm's workspace, our focus centers on data aggregation, particularly prioritizing geospatial data acquisition. This involves the meticulous collection of data, including key indicators such as unemployment rates, employment statistics, job listings, popular Points of Interest (POIs), building permit data, and the influential Zillow Home and Rent index. These diverse inputs form the foundation for our analytical models, driving the investment strategies at the core of the company's operations. Through the careful synthesis of these data sources, we equip our investment decisions with the insights necessary for success.

3.5 Description of Data

3.5.1 RMBS Dashboard Data

The data used to construct the RMBS Dashboard originated from various sources, including Zillow, Redfin, Esri, and the firm's internal Multiple Listing Service (MLS). The raw data obtained from these sources comprised a diverse set of metrics related to both company-specific and broader regional real estate market data. Such metrics included, but were not limited to, county names, home price index, and homes sold.

3.5.2 LLM Deal Appraisal Prompt Data

The data used to develop the LLM Deal Appraisal Prompt originated from the investment presentations created by the company's internal deal desk. These presentations were designed by investment analysts to present to the company’s Investment Committee (IC) and contained a variety of pertinent information such as company name, annual recurring revenue, market statistics, etc. Originally created in a PowerPoint slide format, these presentations were then converted to a Portable Document Format (PDF) for our team to process.

4. Methodology

4.1 Scrum Software Development Methodology

Scrum is a framework that helps software teams “generate value through adaptive solutions for complex problems” (Schwaber & Sutherland, 2020). It embodies the values and principles of Agility and emphasizes the frequent delivery of a product to the client. To complete a product using Scrum, the team determines which software features are important and these become the building blocks of the product known as user stories. User stories are tangible pieces of value that are written from the perspective of the user, and they define the who, what, and why of that feature. These user stories are then put into the product backlog, which is a list of work for the developers (Atlassian, 2019). The next component of Scrum is the Sprint. A sprint is a time-boxed period, typically between one and four weeks long, where a team completes a set amount of work (Atlassian, 2019).

The Scrum Product Management system adheres to three main roles inside of a Scrum team: Product Owner, Scrum Master, and Developer(s). While these roles are pre-set, they do not imply any hierarchy within the team but rather the responsibilities of each team member. The Product Owner’s main role is to define the user stories of the team and create a product backlog. The Scrum Master is responsible for ensuring the team is acting in accordance with the scrum process along with overseeing the mitigation of potential blockers for team members as they arrive. Finally, Developers are team members who are committed to creating any aspect of a usable Increment in each Sprint. While each position has a robust role within the scrum team, all members work collaboratively towards the sprint goal.

Scrum contains five main ceremonies: backlog grooming, sprint planning, daily stand-ups, sprint review, and sprint retrospectives (Gareth, 2021). Backlog grooming is a procedure

where the team will discuss and prioritize user stories in the backlog. It should take between 45 minutes and an hour and can include removing outdated stories and adding new ones. Next for the sprint planning, members of the agile team will decide which user stories to pull from the backlog to include in the upcoming sprint. Then each workday, the team will meet during daily stand-ups, often at the same time and place, and discuss what everyone has been working on and if there are any roadblocks present. Stand-ups are time-boxed to 15 minutes and are organized and facilitated by the scrum master. At the end of a sprint cycle, the team will then have a sprint review and sprint retrospective. The purpose of a sprint review is to go over what was completed in the sprint with the key stakeholders and collect feedback. The final scrum ceremony is the retrospective. During a retrospective, the team goes over what worked well and where there are pain points in their product management methods so that they can make improvements for the next sprint.

4.2 Why Scrum?

Agile methodologies offer numerous advantages compared to other approaches. One of the notable benefits is adaptability even in later stages of development. During sprint reviews, team members present the sprint deliverables to stakeholders and identify potential improvements or adjustments to be made. This frequent engagement with stakeholders allows the team to quickly adapt the product to meet the stakeholders' needs. Scrum also includes a retrospective component that allows the team to determine pain points in their product management practices and make the proper changes to improve productivity for the next sprint. These meetings with stakeholders and retrospectives allow the team to address concerns with their product or methodology and make the appropriate changes promptly.

Another benefit of agile is frequent delivery. The goal for each sprint is to deliver a valuable increment of work to the stakeholders. This is beneficial because customers can receive working pieces quicker than they would if the team used a different methodology. This is especially helpful for companies that benefit from having working parts of a larger product. Frequent delivery also allows the team to take immediate action if the stakeholders' interests are not met or if their interests change.

Scrum also emphasizes collaboration and transparency within the team. Since there are daily stand-ups, the team is more informed on what other members are working on and can stay on top of any roadblocks that occur. By meeting frequently and talking about blockers, members are more transparent and can get quick assistance when needed. Furthermore, Scrum defines distinct roles, each of which plays a crucial part in the overall agile process. The presence of assigned roles minimizes confusion in the workflow, as individuals are responsible for their delegated tasks.

4.3 RMBS Dashboard Methodology

4.3.1 Project Goal & Objectives

The overarching goal of this project was to enhance the company's strategic decision-making process by conducting a comprehensive analysis of its historical deal demographic data. Before our involvement, the company had a version of an RMBS dashboard that contained outdated data. Thus, the company wanted our team to preserve the types of data visualizations shown in our newer iteration while also updating the data ingestion pipelines to increase insight relevancy. The aim was to provide a nuanced understanding of the firm's historical deal landscape, enabling stakeholders to make informed decisions, identify trends, and optimize

future investment strategies based on data-driven insights. To achieve this goal, our team identified the following objectives:

1. Replicate the metrics and visual templates embedded in the company's current Residential Mortgage-Backed Securities (RMBS) deal dashboard and allow them to keep the data updated with Databricks.
2. Enhance the data pipelines and user interface of the dashboard to optimize information retrieval and ensure up-to-date representation of metrics.

4.3.2 Data Overview

Our sponsor supplied the team with data sources from their previous RMBS deal dashboard, incorporating publicly available datasets from real estate companies Zillow and Redfin, along with internal deal database information from their Multiple Listing Service (MLS) desks. Due to their large size, these datasets were hosted on the company's Data Lake, which served as the origin point from where they could be accessed in Databricks.

4.3.3 Data Manipulation and Transformation

In the initial stages, the team had to carefully consider specific fields essential for designing the dashboard while also addressing various challenges associated with data inconsistencies. All data sources were collected and uploaded to the company's Azure Data Lake. Subsequently, the team utilized this data from their local computers to perform essential data cleaning and export tasks within the Databricks environment. The team's data manipulation programming language of choice was Python, due to its compatibility with pandas tools. The pandas library is a "fast, powerful, flexible and easy to use open-source data analysis and manipulation tool, built on top of the Python programming language" (pandas, 2023).

4.4 LLM Deal Appraisal Prompt Methodology

4.4.1 Project Goal and Objectives

The primary goal of this project was to create a system for the company's investment analysts and Investment Committee (IC) to evaluate investment deals using the company's proprietary ChatGPT utility called Samantha. Given that the company's IC receives hundreds of investment profile presentations from analysts, the task of individually examining each profile becomes extremely time-consuming for their team. This project was initiated with the idea of leveraging ChatGPT's natural language processing capabilities to automatically process and rank every document based on its quality as a potential investment deal. To achieve this goal, our team identified the following objectives:

1. Design a mathematical heuristic to evaluate and appraise economic data present in deal profile PDFs.
2. Design a parser for the deal profile PDFs that ensures the preservation of all information in its original context (e.g., keeping metric-value pairs intact)
3. Prompt ChatGPT to extract specific metrics from the PDFs.
4. Input extracted values into the heuristic to derive a score from one to ten, assessing the quality of the deal.

4.4.2 Data Overview

The company supplied our team with investment profile slideshows in PDF format. These documents contained a variety of pertinent information for economic evaluation, including but not limited to ARR, revenue, and gross margin. In addition to being provided to our team for manual examination in a file, these documents were also accessible through Samantha, allowing retrieval via prompts from the user interface.

The primary challenges posed by these documents related to their format and the inconsistencies in data, both qualitative and quantitative, across company profiles. First and foremost, reading PDFs with LLMs proved exceptionally challenging when it came to accurately preserving the logical structure of graphically formatted information. This meant any form of images (graphs, diagrams, etc.) and slide design formatting (indentations, page divisions, etc.) retained limited information once read by the LLM. Additionally, the data included in these company profiles varied, introducing an element of uncertainty to our heuristic model. Since there was a limited number of metrics present across *all* deal files, it was difficult for us to prompt and appraise deals in a way that was consistent while still making economic sense.

4.4.3 Data Extraction

In response to the data challenges, our team needed a parsing program capable of scanning text in accordance with the document's formatting features. Luckily, there is a variety of cloud-based services designed for document data extraction, ranging from specialization in image recognition to regular text. Typically, such programs are easily integrated through import statements within the development environment. By utilizing such services, parsing functions can become considerably more robust in recognizing a greater range of character recognition.

In addition to an advanced parsing program, the team would also need JavaScript Object Notation (JSON) reading and writing functionalities within the coding infrastructure. Unlike alternative forms of file formatting, JSONs offer a better "format for representing structured data based on JavaScript object syntax" (Mozilla, 2023). In essence, JSON files are easier for developers to logically interpret because they preserve complex nested structures within data.

This feature is especially valuable to the team because the deal document data possesses a high degree of complex data relationships.

4.4.4 Heuristic Metric Range Development

ARR Growth Year-over-Year

ARR growth rate benchmarks are highly dependent on company sizes, fiscal year, and current market trends by industry. However, for our project's purposes and limitations, we decided to average these benchmarks into general ranges for review. As specified by MetricHQ, a company dedicated to KPI reporting, "For SaaS companies at 1-5M in ARR, the median YoY ARR Growth Rate is between 52% and 59%, and the top quartile is between 102% and 154%. For SaaS companies at 5-15M in ARR, the median YoY ARR Growth Rate is between 46% and 55%, and the top quartile is between 100% and 131%." (*ARR Growth Rate*, 2023). Using this baseline, our team decided on the following ranges to utilize in the heuristic:

Quality	Benchmark	Score
Poor	< 20%	2
Below Average	20% <= x < 40%	4
Average	40% <= x < 60%	6
Good	60% <= x < 80%	8
Very Good	>= 80%	10

Table 4.1: ARR Growth Year-over-Year Ranges

CAGR

Continued Annual Growth Rate (CAGR) is a financial metric that assists companies in valuing their products and services, as well as evaluating their return on investment. CAGR can

also provide insight into a company's average performance over a specific number of years (Shark Finesse Ltd, 2021). For the purpose of our project, we followed the guideline that "8% CAGR is poor... companies who have been around for 10 or more years may see a CAGR of 8%-12% which is a good rate of sales... Smaller companies should usually aim to see a CAGR of between 10%-20% and start-up businesses may see a much higher rate of growth with numbers as high as 100%" (Shark Finesse Ltd, 2021). As CAGR can fluctuate based on the size of the company, we have decided to create a decision tree for our heuristic to follow, depending on the size of the company.

CAGR Small

Quality	Benchmark	Score
Below Average	0% < x <= 75%	4
Average	75% < x <= 100%	8
Good	x > 100%	10

Table 4.2: CAGR Small Ranges

CAGR Mid-size

Quality	Benchmark	Score
Below Average	0% < x <= 8%	2
Average	8% < x <= 15%	4
Good	15% < x <= 30%	8
Very Good	x > 30%	10

Table 4.3: CAGR Mid-size Ranges

CAGR Large

Quality	Benchmark	Score
Below Average	0% < x <= 3%	2
Average	3% < x <= 5%	4
Good	5% < x <= 12%	8
Very Good	x > 12%	10

Table 4.4: CAGR Large Ranges

CARR Growth Year-over-Year

Contracted Annual Recurring Revenue (CARR) is the baseline for understanding recurring revenue in SaaS companies. This metric quantifies the annual recurring stream of revenue generated from subscription-based contracts. CARR relies solely on the size of the company and the quantity of pre-existing customers they assist. The larger the CARR, the better the company can retain customers (mosaic, 2023). When comparing year-over-year CARR growth, an increase indicates growth for a company. We based these benchmarks using similar ranges to ARR Growth.

Company	Benchmark	Score
Poor	< 15%	2
Below Average	15% <= x < %25	4
Average	25% <= x < 50%	6
Good	50% <= x < 60 %	8
Very Good	x >= 60%	10

*Varies on company size

Table 4.5: CARR Growth Year-over-Year Ranges

EBITDA

The earnings before interest, taxes, depreciation, and amortization (EBITDA) are best used when comparing two companies in the same industry, but they can also be used by investors to assess a company's operational efficiency. A company needs to demonstrate positive growth in its EBITDA because, while it only provides a snapshot of a company's financial health, it can be a good indicator of how well a company is operated (Scott, 2022). As seen in the table below, the range for an average-quality EBITDA is between zero and fifteen million. The range is broad because, as long as the company can generate a profit, it demonstrates a level of stability, making it investable.

Quality	Benchmark	Score
Below Average	< $0	2
Average	$0 < x < $15 million	8
Good	x >= $15 million	10

Table 4.6: EBIDTA Ranges

Gross Margin

Since gross margin is a benchmark to determine a company's ability to generate a profit compared to its revenue, it is a good metric to consider when deciding to invest in a company. Each industry has vastly different gross margins; however, across industries, 30% gross margins are considered the overall average (Bloomenthal, 2022). For this reason, anything below the industry average receives a lower score, as it would be unnecessary to consider deals that do not keep up with the profit generation of the economy.

Quality	Benchmark	Score
Poor	<= 5%	2
Below Average	5% < x <= 10%	4
Average	10% < x <= 20%	6
Good	20% < x <= 30%	8
Very Good	x > 30%	10

Table 4.7: Gross Margin Ranges

Gross Retention

Like ARR growth, gross retention benchmarks also benefit from the inclusion of broader industry metrics. However, lacking the necessary resources to effectively specify this information, our team has decided to generalize using the following scale: “Median GRR is approximately 90% across all Software as a Service (SaaS) companies. For those selling to small and medium businesses, a good GRR is 80%. For Enterprise SaaS, 90% is a good GRR. For very high Annual Contract Value (ACV) products, you should benchmark your GRR up to 95%” (Lido, 2023). Taking these estimates into account, our team decided on the following benchmark values for our heuristic:

Quality	Benchmark	Score
Below Average	< 80%	2
Good	80% < x <= 90%	5
Very Good	> 90%	10

Table 4.8: Gross Retention Ranges

Magic Number

Generally, the Magic Number is a metric used in companies that provide Software-as-a-Service to understand sales efficiency. This value is especially important in gauging the strength and marketing efficiency within a firm. To calculate the Magic Number for a company, divide the annual recurring revenue from the current quarter by the total customer acquisition cost (Mosaic, 2023). Benchmarks below 0.75 indicate that a company is "on the right track with sales efficiency," while benchmarks greater than that show the company can "build out the sales marketing strategy" (Mosaic, 2023).

Quality	Benchmark	Score
Poor	<= .5	2
Below Average	.5 < x <= .75	4
Good	.75 < x <= 1.0	8
Very Good	1.0 < x <= 1.6	10

Table 4.9: Magic Number Ranges

Net Retention

Since net retention represents the overall trend of a business's revenue growth, it is best for this value to be over 100%. Anything below this indicates a company that is actively losing revenue. Thus, for this metric, it is generally accepted that higher values are better, as "A high net revenue retention rate shows predictable and scalable growth. And the higher the rate the better your company's prospects with investors" (Hayes, 2022). Therefore, our team implemented the following scale for this metric:

Quality	Benchmark	Score
Below Average	< 100%	2

Good	= 100%	8
Very Good	> 100%	10

Table 4.10: Net Retention Ranges

Revenue Growth

Revenue growth tracks the year-over-year change in total gross sales. This metric is also a critical driver of corporate performance as it reflects the general success of a business. Thus, the greater the positive change in growth, the more a company has profited from sales. Locating concrete benchmarks for revenue growth performance is considerably difficult, as it is typically evaluated against the growth of a particular company's industry. For our project's purposes, we decided to derive a metric based on the growth of the overall economy, which in the second quarter of 2023 was 2.1% (Bureau of Economic Analysis, 2021). As such, our heuristic judged good revenue growth by whether the company was exceeding the performance of the US economy.

Quality	**Benchmark**	**Score**
Below Average	< 2.1%	2
Average	= 2.1%	5
Good	> 2.1%	10

Table 4.11: Revenue Growth Ranges

TAM

Total Addressable Market, or TAM, targets the estimated demand for a specific product or service. This metric can be used to calculate the implied revenue opportunity for a particular company (Wall Street Prep, 2022). Fluctuating based on the market size and company seniority,

TAM for a startup company can range anywhere between $10 million to $300 million (Wing, 2023).

Quality	Benchmark	Score
Below Average	< $10 million	2
Average	$10 million < x <= $300 million	5
Good	> $300 million	10

Table 4.12: TAM Ranges

TEV/CARR

The TEV/CARR Multiple is an internal company metric that illustrates the total company valuation over the predicted recurring revenue from an individual customer. This value is an indicator of the trend the company is on, with a negative trend indicating growth. The smaller the TEV/CARR multiple, the better, showcasing larger recurring revenue for the company year-over-year.

Company	Benchmark	Score
Below Average	< 5	2
Average	5 < x < 10	5
Good	> 10	10

*Varies on company size

Table 4.13: TEV/CARR Ranges

5. Software Requirements

5.1 Software Requirements Gathering Strategy

Our project requirements were primarily determined by our sponsor, who met with us on a daily cadence. During these meetings, we discussed our progress on the final goal, obstacles encountered during our workday, questions about the product, and the direction of the project.

5.2 Functional and Non-functional Requirements Gathering Strategy

5.2.1 Requirements – RMBS Dashboard

The main requirement of this project was to create a dashboard intended to demonstrate the company's historical rate mortgage-backed security data. Viewing this data is important because it demonstrates the performance of the firms' investments across the United States. In addition, the requirements for this task specified that our team had to pull and transform data from the following sources: ESRI, Redfin, Zillow, New Home Source, and Multiple Listing Enterprise Solutions CoreLogic. Once pulled, this data must then be validated to make sure the correct information is being used to create the dashboards. Moreover, once the data was validated, a pipeline using Azure Databricks and PostgreSQL must be created to transfer the data into Power BI, the company's dashboarding software. The final requirement involved creating a filter to select a deal that would update all the data shown on the dashboard.

5.2.2 Requirements – LLM Deal Appraisal Prompt

The large-scale functional requirements for the LLM deal appraisal prompt deliverable were to develop a model capable of generating an appraisal value (ranging from one to ten) for the company's investment profiles using the Samantha chatbot. By providing a swift metric to

assess investment quality, this project aimed to determine the viability of hundreds of deal documents to save the Investment Committee's time by facilitating the straightforward identification of investment profiles worthy of manual review. Additionally, this project required us to find new ways to extract data from PDF's and presentations while maintaining proper context. Analyzing the content within these investment profiles enabled our team to consider various financial metrics as a basis for calculating the profile scoring.

5.3 User Stories and Epics

Sprint	User Story	Points
Epic: RMBS Tear Sheet		
1	As a stakeholder, I want to view redfin data with certain conditions so that I can better understand the data and its fields	5
1	As a stakeholder, I want to view demographics data with certain conditions so that I can better understand the data and its fields	5
1	As a stakeholder, I want to view new home source data with certain conditions so that I can better understand the data and its fields	5
1	As a stakeholder, I would like to see a Power BI dashboard that replicates an existing dashboard so that I can easily see the data	8
2	As a stakeholder, I would like to have a filterable view of organized demographic data [tear sheet 2-4] so that I can view demographic date for a particular deal	3
2	As a stakeholder, I would like Zillow and Redfin data to be connected to deal data in Data Lake so that I can filter various metrics by deal	5
2	As a stakeholder, I would like to have a table view and plot views of redfin data filterable by deal [tear sheet 9-15] so I can see the data for a specific deal	8
2	As a stakeholder, I would like to have filterable views of the weekly redfin data that display multiple regions in their plots so I can view the data for different deals [31-34]	8

2	As a stakeholder, I would like to have filterable views of Zillow data [17-19] so I can view the data by deal	5
2	As a stakeholder, I would like to have a filterable table view of new home source data so I can understand what new homes around a surrounding deal look like	5
2	As a stakeholder, I would like to have filterable views of MLS data [24-28] so I can view the data by deal	5
2	Carry over the Power BIs and Code from previous week to VM	2
3	Submit final Power BI for approval	5
Epic: LLM Deals		
3	Practice accessing company's prompt engineering code with Samantha	8
3	Review the parent company's deal documents	3
3	Meet with Prof Blais to discuss heuristic	5
4	Meet with Prof Sarnie to discuss heuristic	2
4	Preliminary testing scoring ICs without heuristic	2
4	Meet with Professor Dunbar to develop a heuristic	2
4	Develop a list of heuristics for ICs	5
4	implement data bricks index operations	8
4	Develop the prompts in VS Code to understand how the document is being read.	13
5	Put metrics into JSON object and print out to verify	3
6	Read in metrics from JSON object and apply some inequality operations (no need to be final ranges, but just do some for proof of concept) then output a final score	8
6	Implement a way to account for metrics that do not appear in a document (either cut them in final calculation or prioritize)	2
6	Pass in the formatted tables into the index	3
6	Determine ranges for metrics we are using	8
6	Refine the list of which metrics we are using	2
7	Run model several times to ensure consistency	3
7	Add a project field into the Samantha output and the PostgreSQL dataframe	1
7	Alter output for better readability	3
7	Implement ranges for CAGR considering company size	3
7	Update YoY CARR method to fit in the range 1-10	1
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 5.3: Completed User Stories from Development Epics

6. Design

Our Major Qualifying Project was comprised of two distinct subprojects with different functionalities. Consequently, we navigated two distinct workflows and architectures while handling these assignments.

6.1 High Level Architecture: RMBS Dashboard

6.1.1 Technology Framework

The RMBS Dashboard project utilized the company's Azure web services, including Azure Data Lake and Databricks in addition to Python, pandas, PySpark, PostgreSQL, DBeaver, and Power BI. The overarching objective of our project was to intake RMBS data from both Data Lake and PostgreSQL, subsequently recreating dashboards with automated updating capabilities.

Before hosting the data in Data Lake, we leveraged Databricks to read company deal/real estate data from various sources. Our process involved utilizing PySpark, Python, and the pandas library to establish a functional ETL pipeline. For data that did not require further transformation, we opted to host it on PostgreSQL. This data was then directly integrated into Power BI, as illustrated in Figure 6.1.1.

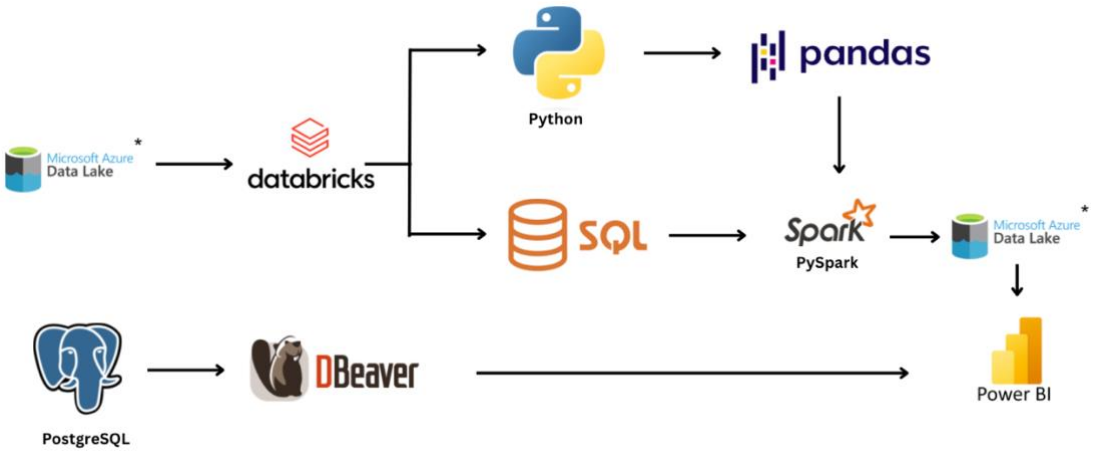


Figure 6.1.1: Project 1's Architectural Design Flowchart

Figure 6.1.1 shows an abstract view of the technologies used in this initial project and the workflow behind them. Overall, Azure Databricks and its notebook interface was our primary location for writing PySpark and Python code. With Python and the pandas library, our team was able to create dataframes which proved to be a more efficient way to query and extract necessary data for each deal. Since PySpark utilizes Python, we were then able to pull these dataframes directly into PySpark and write out tables into Azure Data Lake. These Azure web services allowed us to deploy an automated process to extract, transform, and load data into Power BI to be seen through dashboards.

6.1.2 Entity Relationship Diagrams

Entity Relation Diagrams (ERDs) serve multiple purposes in database management and design. These diagrams demonstrate how tables of data are normalized to reduce redundancies in the overall schema. Ensuring proper relationships increases data integrity and, in turn, allows for more consistency across the board. In our team's case, we used a many-to-many relationship within the 'deal' column across all tables. Leveraging the Slicer feature in Power BI, we

seamlessly connected all tables throughout the dashboard. This enabled the application of a unified filter, ensuring that any adjustments made would dynamically update all associated tables across the entire sheet.

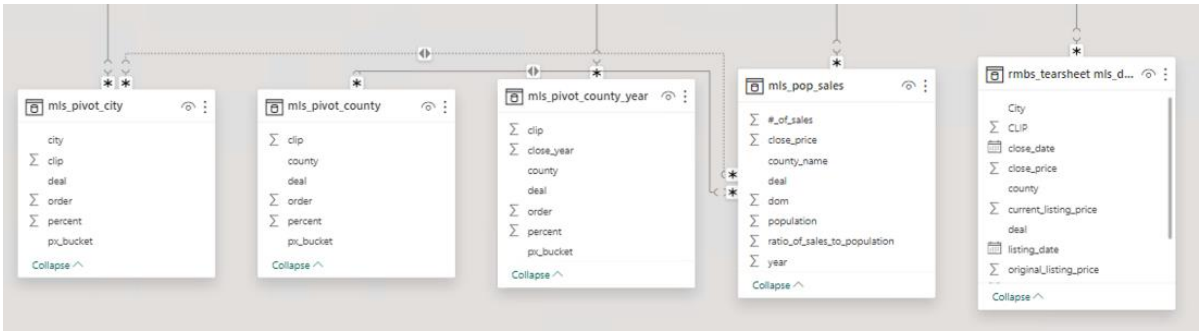


Figure 6.1.2: Entity Relationship Diagram for MLS Data

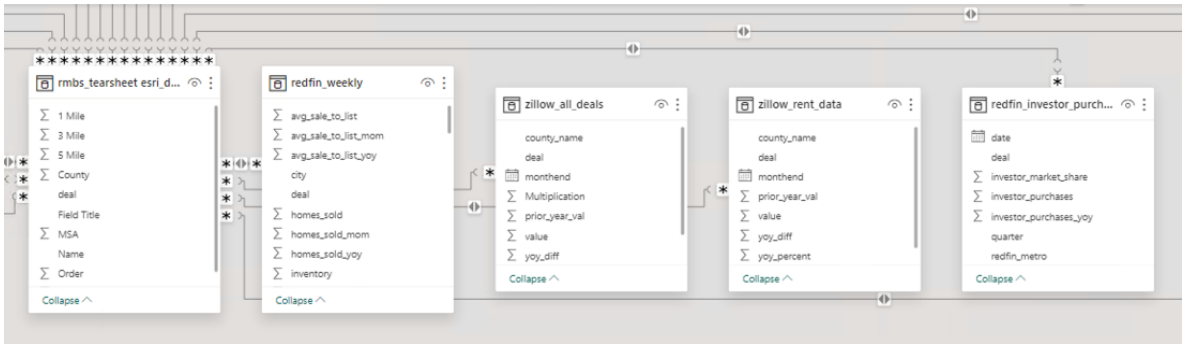


Figure 6.1.3: Entity Relationship Diagram for Redfin, Zillow, and ESRI Demographics Data

As seen in Figure 6.1.2 and Figure 6.1.3, the Power BI dashboard corresponded to multiple different tables. This data was uploaded through the Databricks cloud Java Database Connectivity application to allow for the ease of updating data in the Power BI dashboard. These tables were all connected through normalized relationships to ensure consistency throughout the entire dashboard. In total, there are 22 tables consisting of different data from varying data sources that are connected through the ‘deal’ column in the RMBS tear sheet.

6.1.3 Dashboard Design

The dashboard's design is structured based on the distinct sources of data it encompasses, such as grouping Zillow and Redfin data together. Drawing inspiration from the prior tear sheet, the sheet's style was curated. The dashboard is split up into four separate pages: ESRI Demographics, Redfin and Zillow Single Home Residential, Redfin New Home, and Multiple Listing Service data. Through an array of structured tables, bar graphs, and line graphs, the data is presented in an organized manner. To ensure a user-friendly experience, we implemented the slicer feature in Power BI, leveraging a many-to-many relationship. This strategic choice facilitates seamless data filtering, contributing to a more intuitive and efficient user interface.

6.2 High Level Architecture: LLM Deal Appraisal Prompt

6.2.1 Technology Framework

In the team’s second project, a different technology stack was required. In this phase, we were given deal sheets sourced from the Investment Committee at the parent company. To thoroughly evaluate the intricacies of each deal, we leveraged the alternative investment firm’s proprietary ChatGPT interface, enabling us to create appraisals tailored to the unique aspects of each specific deal.

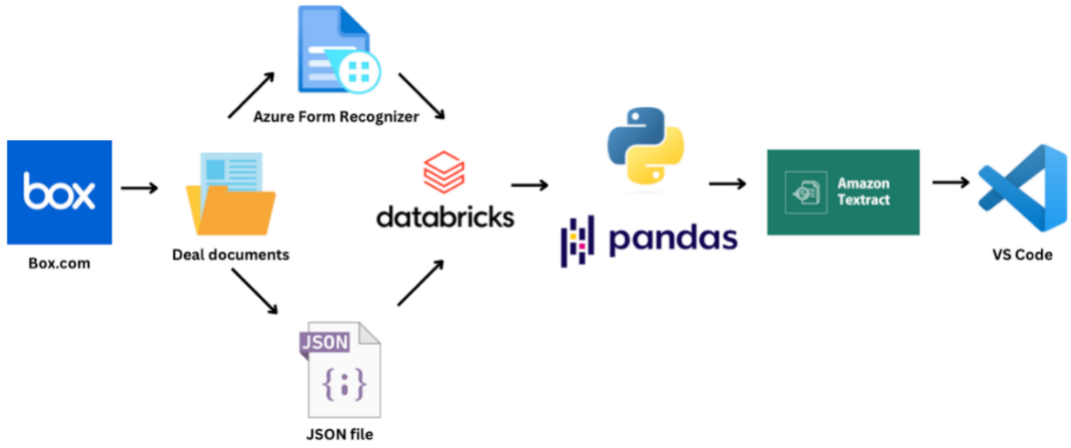


Figure 6.2.1: Project 2's Architectural Design Flowchart

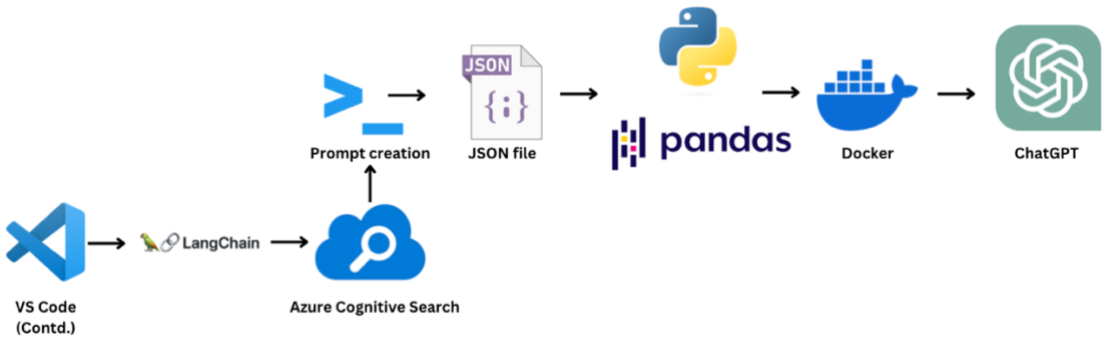


Figure 6.2.2: Project 2's Architectural Design Flowchart (Contd.)

Figures 6.2.1 and 6.2.2 present a high-level overview of the various steps taken to reach our final working model. This process, illustrated in the two figures, demonstrates the flow of information procedures. Upon obtaining the deal sheets from Box.com, a secure cloud document management webpage, and converting them into JSON files, the developers utilized Databricks and Python to commence the initial step in uniformly parsing the data. In this process, an index was created to facilitate further data transformation using models in VS Code. Following this, we utilized LangChain, a Python library, to read the index and apply cognitive search to the JSON

file. The result was then transformed into a dataframe, marking the conclusion of the data transformation process. After successfully transforming the data, the team proceeded to create a prompt instructing the LLM on what parameters to search for and how to quantify the category based on the deal data. This approach facilitated a more targeted and effective utilization of the LLM in the subsequent stages. The final step involved testing the prompt and assessing how the Language Model (LLM) ingests information, accomplished through the utilization of Docker containers to create an unpublished version of the model.

7. Implementation

Our team used Agile Scrum methodology throughout the project. We chose to use an Agile Scrum approach as opposed to other product management methodologies because Agile welcomes changing requirements and allows for frequent software delivery. Daily standups were performed each day at 10:00 AM EST, lasting 15 minutes, and were facilitated by the team's scrum master. We invoked one-week sprints, with each sprint starting and ending on Tuesday. Our sprints were tracked using Jira Software. Jira has a sprint board that displays all the current tasks that the team is working on in that sprint as well as a product backlog which is an organized list of user stories that are yet to be completed and are not in a current sprint.

At the start of every sprint, the team created user stories and each group member assigned a story to themselves when they began work on that task. The created stories directed the team to what should be done during that week, however, could be modified as seen necessary. At the end of each sprint were sprint reviews which occurred weekly on Tuesdays at 10:30 AM with the entire team and our sponsor. Following that were sprint retrospectives conducted in the project management tool, Trello. During sprint retrospectives, the scrum team discussed how the past sprint went, how to improve, and techniques to do scrum more effectively. The times and lengths of the events were determined by what is traditional for Agile Scrum as well as the availability of the sponsoring employees. It was important to our team to use one-week sprints as the entire length of the project was seven weeks and we wanted to make sure we presented deliverables frequently so we could reveal problems quickly, thus allowing us to make any necessary changes in a timely manner. This tempo allowed us to build out our sprint backlogs and ensured we were working efficiently during our sprints.

7.1 Weekly Sprints

7.1.1 Sprint One

For our first sprint, our team met with the main point of contact within the company, to discuss the expected deliverables for the week. Given that we did not have access to the company system, the liaison was able to create an offshore server to host some data we would be using. In this server, we were given a database with relations of locational housing data from various sources like Zillow and Redfin. The liaison also gave us a Python script to access the database and clean the data that we query along with other functionalities. Additionally, we were given a PDF that had tables and plots created in Power BI. Our goal for the week was to replicate the PDF by querying the data in the server and pulling the necessary 'deals' data from the appropriate schemas. Since half the group did not have access to Power BI, some members took further responsibility with the report along with using Python to connect the central database to PostgreSQL. An additional goal was to submit the paper for further review with modified Background and Methodology which was completed earlier in the sprint. Below are the user stories from Sprint 1 with a column specifying if they have been completed.

Completion Confirmation	User Story	Points
	Epic: RMBS Tear Sheet	
Yes	As a stakeholder, I want to view redfin data with certain conditions so that I can better understand the data and its fields	5
Yes	As a stakeholder, I want to view demographics data with certain conditions so that I can better understand the data and its fields	5
Yes	As a stakeholder, I want to view new home source data with certain conditions so that I can better understand the data and its fields	5
No	As a stakeholder, I want to view MLS page data with certain conditions so that I can better understand the data and its fields	5
Yes	As a stakeholder, I would like to see a Power BI dashboard that replicates an existing dashboard so that I can easily see the data	8

No	As a stakeholder, I would like Zillow and Redfin data to be connected to deal data so that I can filter various metrics by deal	5
	Epic: MQP Paper	
Yes	Add agile background to the methodology section	3
No	Create a rough draft for the literature review	8
Yes	Begin the acknowledgement section	1
No	Create a rough draft of similar applications in the background	6
No	As an employee, I would like to watch the RMBS documentary, to better understand mortgage-backed securities	1
Yes	Decide on a citation manager	3
	Epic: Onboarding	
Yes	As a new employee, I would like to be properly onboarded so that I can use the company's resources and access the appropriate data	5
	Total Points Completed	35/60
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.1: Sprint One User Stories

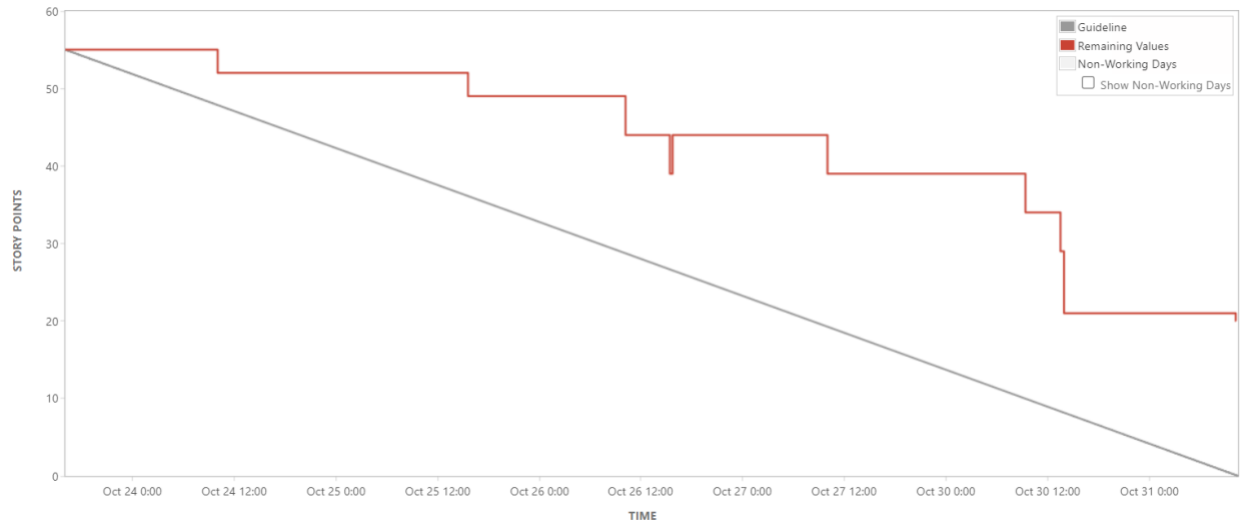


Figure 7.1: Sprint One Burndown Chart

During this sprint, the team successfully completed the following: submitted the paper for review, connected central deal data to PostgreSQL, established clear contact with the sponsor, set up discord chat to share materials, practiced with PostgreSQL, Power BI, and DBeaver and successfully replicated key tables in Power BI. However, during this sprint, the team also

encountered challenges such as missing access to the company's remote desktop and relevant services, initial miscommunication over sponsor tasks and goals, and difficulties with local software access.

Our sprint concluded with a sprint retrospective conducted on October 30, 2023, covering the period from October 23rd to October 27th. The retrospective involved key team members, including Dante Amicarella (developer), Sarah LaRusso (scrum master & developer), Maya Liao (product owner & developer), and Nathan Shemesh (developer), to reflect on Sprint 1. It lasted 30 minutes, and the team discussed what went well, what did not go well, and ideas for improvement.

Some topics that went well included daily meetings with our company liaison, meeting with our Scrum Advisor, making considerable progress on the paper, completing SQL queries and Power BI plots, and overall feeling more comfortable with the tools used. What did not go well included not yet having company accounts and miscommunication about sprint delivery goals with our sponsor earlier in the sprint. Additionally, a few team members expressed feelings of early burnout.

Overall, in Sprint 1 the team excelled in staying focused on tasks through meeting times, clear and respectful communication, attendance, and division of labor. The primary areas that could be improved in future sprints relate to sponsor communication for tasks. We decided to ask for more example walk-throughs to ensure clarity early in the sprint. We also intended to establish official access to the company's system early in the coming sprint and to implement the changes discussed during the retrospective as we saw necessary. To address feelings of burnout, the team brainstormed ideas including changing some days to hybrid or remote days, making sure to take minibreaks, and having a team dinner to socialize in a non-work setting.

7.1.2 Sprint Two

At the beginning of sprint two, we were given access to our accounts so the first day was primarily setting up and understanding where we would be pulling data from. While some data existed on a Postgres database, others were in Azure Databricks and needed additional cleaning, merging, and extra columns. Some of our stories revolved around these tasks. Our overall goal for the week was to continue replicating the existing RMBS tear sheet by using data from the firm's data sources (Postgres and Databricks) and adding filters to see data for a particular deal when they were selected. Associated user stories can be seen below.

Completion Confirmation	User Story	Points
	Epic: RMBS Tear Sheet	
Yes	As a stakeholder, I would like to have a filterable view of organized demographic data [tear sheet 2-4] so that I can view demographic date for a particular deal	3
Yes	As a stakeholder, I would like Zillow and Redfin data to be connected to deal data in Data Lake so that I can filter various metrics by deal	5
Yes	As a stakeholder, I would like to have a table view and plot views of redfin data filterable by deal [tear sheet 9-15] so I can see the data for a specific deal	8
Yes	As a stakeholder, I would like to have filterable views of the weekly redfin data that display multiple regions in their plots so I can view the data for different deals [31-34]	8
Yes	As a stakeholder, I would like to have filterable views of Zillow data [17-19] so I can view the data by deal	5
Yes	As a stakeholder, I would like to have a filterable table view of new home source data so I can understand what new homes around a surrounding deal look like	5
Yes	As a stakeholder, I would like to have filterable views of MLS data [24-28] so I can view the data by deal	5
Yes	Carry over the Power BIs and Code from previous week to VM	2
	Epic: MQP Paper	
Yes	As a reader, I would like to see an updated software development section for sprint 1 (chap 7) so that I could understand the team's methods and any changes to their agile from retrospective	3

Yes	As an employee, I would like to watch the RMBS documentary, to better understand mortgage-backed securities	1
Yes	Update software requirements chapter of the paper as needed	2
	Epic: Onboarding	
Yes	Finish Databricks Tutorial	1
Yes	As a new employee, I would like to connect to all needed parts of tech stack so that I can utilize these tools for my project	5
	Total Points Completed	53/53
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.2: Sprint Two User Stories

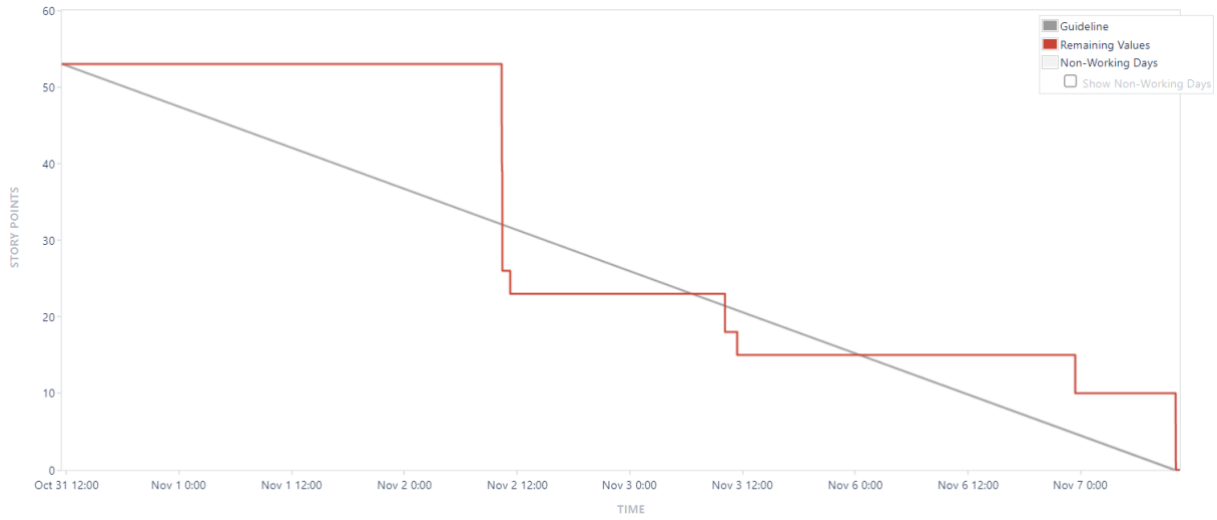


Figure 7.2: Sprint Two Burndown Chart

During the sprint, the team successfully connected Power BI to the appropriate data sources, created additional methods to add edited tables to Databricks, and completed all the dashboards that we set out to; thus, finishing the RMBS tear sheet epic. The team, however, did not complete all the stories for the MQP paper as they were put more on the back burner to prioritize product work.

The sprint retrospective was conducted on November 7, 2023, covering the period from October 31st to November 6th. The retrospective involved key team members, including Dante

Amicarella (developer), Sarah LaRusso (scrum master & developer), Maya Liao (product owner & developer), and Nathan Shemesh (developer), to reflect on Sprint 2. The team continued daily meetings with our company liaison which allowed the team to stay on track and ensure that we progressed to our goals and addressed blockers as soon as possible. Additionally, Nathan led the team in compiling all data and charts into one dashboard. What did not go well involved the lack of collaboration when our team was fully remote on Friday. This created difficulties when we tried to merge all the Power BI as it was difficult for us to describe the issues we were having over text. As a result of a lackluster Friday, our team has decided to modify the original structure of all-remote Fridays to a half-remote day.

7.1.3 Sprint Three

During Sprint Three, our team's primary focus was twofold: kickstarting the development of heuristics for assigned categories and advancing the final downtime report. We delved into the development environments, learning, and experimenting as we began researching heuristics. Most user stories centered on heuristic development, while progress continued expanding the report's background. This sprint sets the stage for refining heuristics and fortifying the report in subsequent phases.

Completion Confirmation	User Story	Points
	Epic: LLM DEALS	
No	Collect information on the parent company’s deal ranking heuristics	1
No	Develop heuristic for ICs	5
No	Develop heuristic for CIMs	5
No	Develop heuristic for QLDP private	5

No	Develop heuristic for QLDP public	5
Yes	Practice accessing prompt engineering code with Samantha	8
Yes	Review the parent company's deal documents	3
Yes	Meet with Prof Blais to discuss heuristic	5
	Epic: MQP Paper	
Yes	Research prompt engineering and put it in the background	3
Yes	Research LLMs and put it in the background	3
Yes	Research ChatGPT 4 and include it in the background	3
Yes	Research the parent company and include in background	3
	Epic: RMBS Tear Sheet	
Yes	Submit final Power BI for approval by company liaison	5
	Epic: Onboarding	
Yes	Install docker on the remote desktops	2
	Total Points Completed	35/56
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.3: Sprint Three User Stories

Figure 7.3: Sprint Three Burndown Chart

This sprint opened our eyes to the magnitude of our project and the amount of information we were uncertain about. The team made considerable progress on the final report and understanding the platform where our final deliverable will be tested. On the other hand, the team also fell behind when it came time to create a heuristic for the project. Overall, our team

missed most of the story points since they related to the creation of the heuristics for each category.

The sprint retrospective was conducted on November 13, 2023, covering the period from November 7th to November 12th and involved all team members to reflect on Sprint 3. Many parts went well during the sprint including creating a docker instance to access the company's version of ChatGPT, completing chunks of the paper, and experimenting with half in-person days on Thursday and Friday. The biggest challenge was progress with developing a heuristic for each of the deal document categories as finding potential heuristics was deemed harder than initially expected. To combat this, some ideas the group brainstormed included narrowing our focus to one of the four document categories to determine a heuristic for just one initially and setting up more meetings with professors and students who have more expertise in financial investments.

7.1.4 Sprint Four

During Sprint Four, our primary focus was researching and developing heuristic elements for Samantha in response to our prompt. Through research and numerous meetings with Subject Matter Experts (SMEs) like Prof. Dunbar, we successfully crafted categories enabling Samantha to streamline her document searches. Like the preceding sprint, our user stories revolved around enhancing our heuristic and progressing with our written report.

Completion Confirmation	User Story	Points
	Epic: LLMDEALS	
Yes	Meet with Prof Sarnie	2
Yes	Preliminary testing scoring ICs without heuristic	2
Yes	meet with Professor Dunbar to develop a heuristic	2

Yes	Develop a list of heuristics for ICs	5
No	implement data bricks index operations	8
Yes	Develop the prompts in VS Code to understand how the document is being read.	13
	Epic: MQP Paper	
No	Write Architecture section and create charts at Chapter 6	5
No	Do findings section for first project. Include images from the dashboard (Chapter 8.1)	5
No	Write Chapter 4 (Software Development Environment) of the paper	3
No	Finish the Acknowledgements section of the papers	2
No	Complete the bibliography and check the in-text citations	5
	Total Points Completed	34/52
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.4: Sprint Four User Stories

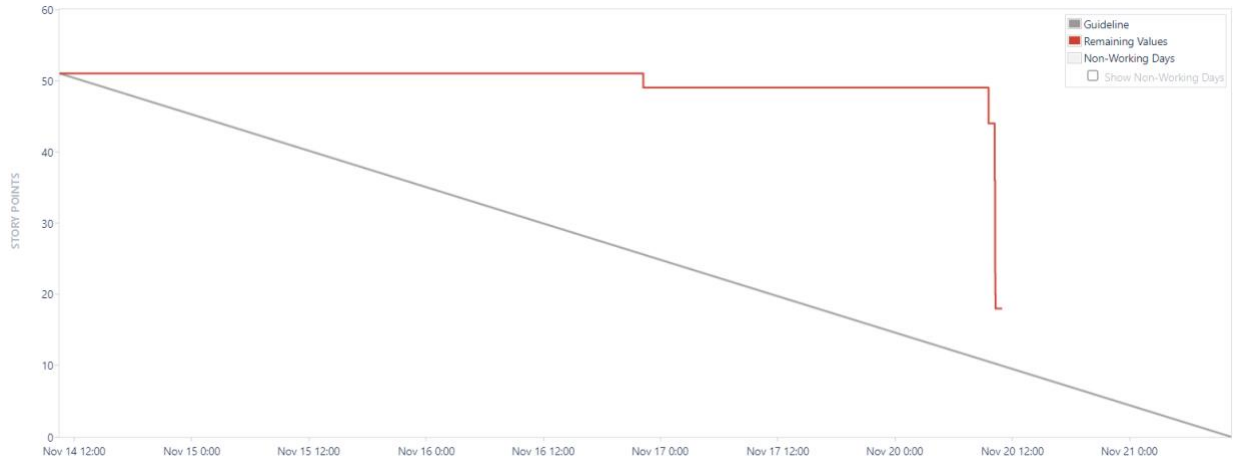


Figure 7.4: Sprint Four Burndown Chart

Overall, this sprint focused heavily on the elements we were going to highlight within our prompt that will be fed to Samantha. The team made a small amount of progress on the report because the main goal was to read into the heuristic elements. Nevertheless, the team finalized the categories going into Samantha, but more is needed to create numerical boundaries for each category. In retrospect, our team felt that Jira did not mirror the amount of effort put into this week and were looking forward to expanding and growing on it.

The sprint retrospective was conducted on November 20, 2023, covering the period from November 14th to November 19th, and involved all team members to reflect on Sprint 4. Following our meeting with Professor Dunbar, we successfully refined the heuristics we were employing. Additionally, we made progress in developing the initial prompt, generating desired metrics in JSON format. However, challenges surfaced, including ambiguity in heuristic value ranges and the need to integrate the JSON file back into our prompt. While we advanced the paper, some sections remain incomplete. To address these issues, we aim to collaborate with our company liaison to streamline our code, allowing us to focus on testing various prompts and dedicating time to finalizing the paper.

7.1.5 Sprint Five

During Sprint Five, our primary focus was working towards a finalized version of our report, concurrently addressing the need to effectively chunk data for proper ingestion to the LLM. At the start of our sprint, we received a new parsing method crafted by our sponsor. Our technical work involved solving bugs created by the new parsing method and creating tables from the parsed data to provide Samantha with an easy-to-understand structure for numerical data. The majority of our stories, however, focused on specific chapter work in our report and ensuring all citations were being marked down.

Completion Confirmation	User Story	Points
	Epic: LLMDEALS	
Yes	Put metrics into JSON object and print out to verify	3
No	Determine ranges for metrics we are using	8
No	Continue narrowing down which metrics to use if necessary	2
No	Implement a way to account for metrics that do not appear in a document (either cut them in final calculation or prioritize)	2
No	Read in metrics from JSON object and apply some inequality operations (do not need to be final ranges, but just do some for proof of concept) then output a final score	8

	Epic: MQP Paper	
Yes	Write and complete charts for Architecture chapter (chapter 6)	5
Yes	Complete bibliography and in-text citations	5
Yes	Add financial metrics and math into background (chapter 2)	3
Yes	Write assessment chapter	5
Yes	Write software requirements chapter (chapter 5)	8
Yes	Add sprint 4 to chapter 7	1
Yes	Complete findings for the first project (chapter 8.1)	5
Yes	Write business and risk chapter (chapter 10)	5
Yes	Write a rough draft for executive summary	5
Yes	Complete chapter 4 (explaining data methods)	3
Yes	Write abstract	3
Yes	Write future work (chapter 11)	5
Yes	Submit rough draft to professors	0
	Total Points Completed	56/76
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.5: Sprint Five User Stories

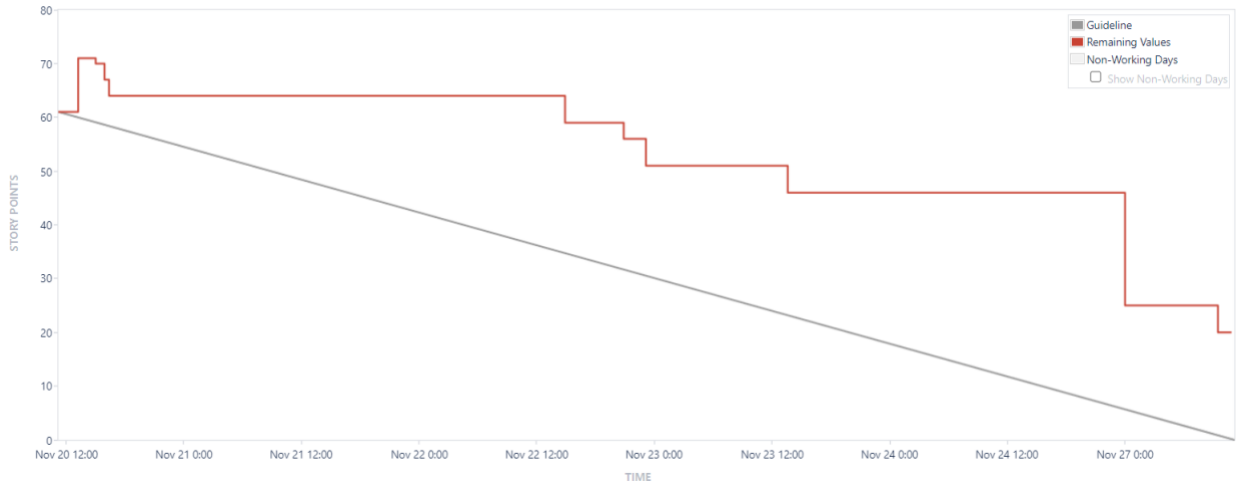


Figure 7.5: Sprint Five Burndown Chart

Overall, this sprint focused heavily on the report, and we completed all the tasks that we set out for it; however, we ran into some blockers for the second part of the project as changing the parsing method necessitated a change to the LLM output format. We had to adjust the chunking size to create chunks of size smaller than the token limit, which caused the LLM to

produce different JSON files, and we still have not loaded the formatted tables into the index. We planned to address these problems during the next sprint through collaboration with our liaison. For the coming sprint, we set out a plan to finish all our development work so that we can focus on our paper for our sponsor to review.

The sprint retrospective was conducted on November 27, 2023, covering the period from November 20th to November 26th, and involved all team members to reflect on Sprint 5. At the beginning of this week, our team took the time to reflect on our CATME assignment and had an open-ended conversation to address how each of us can improve as teammates. Nevertheless, this sprint was undermined by a short week because of Thanksgiving break. As a result, each team member worked remotely and communicated through group messaging, email, Teams, and Slack. When tasks were needed to get done within the paper, team members were able to assign tasks to each other through Word. These concrete lines of communication enabled the team to stay on top of their work and allowed for constant collaboration.

7.1.6 Sprint Six

With it being one of our final weeks of the project, we built out our sprint with high hopes of completing many story points. Our goal for the sprint was to have a working initial version of our deal appraisal running in Samantha that chooses a rating for each of the deals using the metrics we determined earlier by scoring each numerical range the metric could fall in. Another part of our goal was to complete another draft of our paper so that we could send it to the company in the coming sprint.

Completion Confirmation	User Story	Points
	Epic: LLMDEALS	

Yes	Test heuristic implementation to assert consistent scoring of projects	5
Yes	Determine ranges for metrics we are using	8
Yes	Continue narrowing down which metrics to use if necessary	2
Yes	Implement a way to account for metrics that do not appear in a document (either cut them in final calculation or prioritize)	2
Yes	Read in metrics from JSON object and apply some inequality operations (do not need to be final ranges, but just do some for proof of concept) then output a final score	8
Yes	Have Samantha print out a project name (or include this field in the Postgres database)	1
Yes	Pass in formatted tables into the index file	3
	Epic: MQP Paper	
No	Update sections related to LLM Deal Appraisal Prompt project	8
No	Implement rough draft changes based on advisor feedback	8
Yes	Editor for chapter 8	3
Yes	Editor for chapter 3	3
No	Editor for chapter 2	5
Yes	Redact company name from report	5
Yes	Editor for chapter 5	3
No	Editor for chapter 6	3
Yes	Editor for chapter 7	0
Yes	Editor for chapter 10 (business + risk)	3
Yes	Editor for executive summary	3
Yes	Editor for abstract and introduction	3
Yes	Write conclusion of report	5
	Total Points Completed	57/81
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.6: Sprint Six User Stories

Figure 7.6: Sprint Six Burndown Chart

For production work, we created an array of methods in our main Python file in VS Code. These methods take in a metric and return the score for that metric. Afterwards, we considered all the ones that were given, found the average score, and outputted the metric values as well as the final score in Samantha. Additionally, we edited many sections of the paper, implemented most of the changes given to us from advisor feedback, and wrote up sections that were missing from the previous draft.

The sprint retrospective was conducted on December 4, 2023, covering the period from November 27th to December 3rd, and involved all team members to reflect on Sprint 6. For this Sprint, our team primarily focused on writing our report and final development. Throughout this process our team made sure all parts of the report were being written and reviewed by two different team members. Additionally, Nathan and Sarah were able to continue their collaboration efforts in the final touches of the development. On Tuesday, we decided to take some time to go out and eat lunch together. This casual lunch, along with other bonding time, has allowed us to get to know one another outside of the working environment throughout the term.

7.1.7 Sprint Seven

With this being the last week of the project, we finished the final edits for the report and created and practiced the final presentation. We submitted our report and got a redacted version of it back from our company liaison and are finalizing the process of completing the final edits to submit it to the eCDR.

Completion Confirmation	User Story	Points
	Epic: LLMDEALS	
No	Determine how to have ARR not go in CARR	2
Yes	Run model several time to ensure consistency	3
Yes	Alter output for better readability	3
Yes	Add in a project field in the	1
Yes	Implement method for parsing CARG	3
Yes	Fix YoY CARR method	1
	Epic: MQP Paper	
Yes	Remove metrics we ended up not using	1
Yes	Add page numbers for figures, tables, and equations	1
Yes	Editor for chapter 6	3
Yes	Finish final presentation	8
Yes	Update sections related to LLM Deal Appraisal	8
Yes	Add table of equations	1
Yes	Implement rough draft changes based on advisor feedback	8
Yes	Add metric ranges into paper along with its sources	8
Yes	Editor for conclusion	3
Yes	Submit rough draft to AG for review	1
Yes	Editor for background (chapter 2)	5
Yes	Remove figure X in all occurrences	1
Yes	Add math to LLM section in the background	5
Yes	Write results section for LLM Deal Appraisal (Chapter 8.2)	5
	Total Points Completed	69/71
*Red point numbers are incomplete tasks; black numbers are completed tasks		

Table 7.6: Sprint Six User Stories

Figure 7.7: Sprint Seven Burndown Chart

This past sprint, we completed the final presentation, edited, and added subsections of our final report, and performed an analysis on the resulting scores output by Samantha. These final steps were the finishing touches of our project. Additionally, we were able to add two new metrics, CAGR and CARR growth, which allowed Samantha's output to be more precise. Finally, we were able to reformat the final output of Samantha and we also created columns for each metric's score in the PostgreSQL table for easier analysis.

The sprint retrospective was conducted on December 11th, 2023, covering the period from December 4th to December 10th, and involved all team members to reflect on Sprint 7. While this week did not require as much production work to get completed, our team was able to encourage one another to keep plugging away at the final details in our projects. Maya and Dante spearheaded the creation of the presentation while Sarah and Nathan collaborated on the final development of our primary project. Our subgroups were able to cohesively complete the tasks assigned to them.

7.1.8 Sprint Overview

We were able to analyze our agile process over the main seven weekly sprints for our project. Figure 7.8.1 shows a velocity chart of our progress where the number of story points we tasked ourselves with is shown in gray and the number we actually completed is shown in green. While labeled sprints 5-11, these correspond to sprints 1-7 as the first four sprints in our Jira were for preparation and were not discussed in this report. The sprint where we got the greatest number of points completed was the final sprint shown by the largest green bar, and the sprint where we approximated the amount of work to get done the best would be the second sprint as

the gray and green bars are the same height, indicating we completed all of the stories in that sprint.

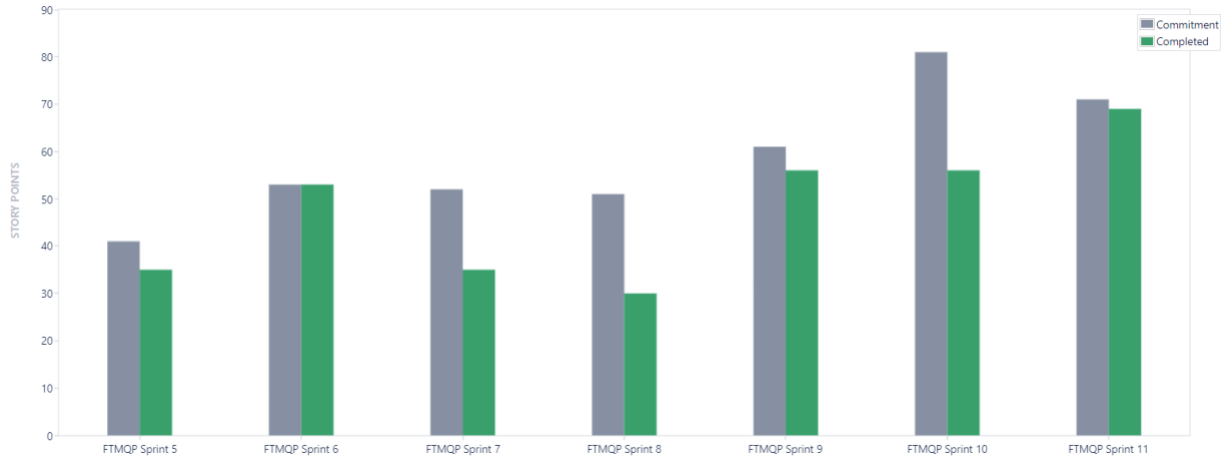


Figure 7.8.1: Velocity Chart for Sprints 1-7

We also analyzed our sprints using a cumulative flow diagram spanning the course of the project. Orange shows what needs to be done, turquoise is in progress, and purple is completed. The y-axis shows the total number of tasks, rather than tracking story points. Where purple is close to orange, this shows where almost all the tasks are completed, while orange far from purple shows when there are a larger number of tasks to be done.

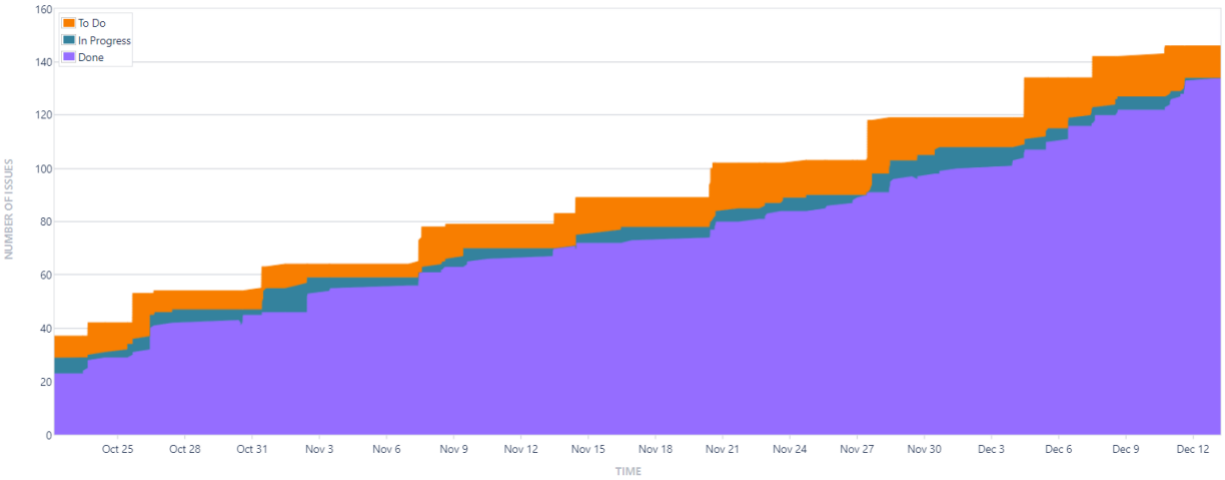


Figure 7.8.2: Cumulative Flow Diagram for Sprints 1-7

7.2 RMBS Dashboard Development

7.2.1 Data Cleaning

The data cleaning process involved making alterations to the company's existing data pipeline to align with the fields specified by the RMBS Dashboard. This encompassed tasks such as removing query exceptions to avoid compiling errors, reformatting column string formatting, and generating new pandas dataframes on the company's Azure Databricks server using Apache Spark. These changes were implemented to facilitate the combination of information, remove unnecessary formatting, and ensure that all column titles were clear and fitting for easy integration into Azure Data Lake.

Following these operations, the resulting dataframes encompassed information related to real estate deal names, year-over-year differences of historical investments, new construction indicators, inventory value, region, and more. The new dataframes were then hosted in Azure Data Lake, and the team imported the information into Microsoft Power BI to make these new data sources accessible for further analysis, enabling efficient data visualization and dashboarding tasks in the future.

7.2.2 Power BI Dashboarding

The final Power BI dashboard hosts a many-to-many relationship on all deal name columns. This relationship across all tables allows for a cascading filter throughout the entire dashboard so that once a deal or multiple deals are selected, only that data is shown. For bar charts, we were able to create bins to group data within certain quantitative ranges. This allowed

the data to be shown in different categories unlisted in the raw data. Finally, we renamed columns that were either unclear or misguiding to the user to provide a cleaner and succinct dashboard.

7.3 LLM Deal Appraisal Prompt Development

7.3.1 Document Data Processing

In response to the data challenges, our team implemented a parsing program capable of scanning text in accordance with the deal documents' formatting features. For our project's purposes, the parser's selection was limited to the first five pages of each document since these pages contained most deal summary financials. This parsing function was developed using Azure's Form Recognizer and AI Document Intelligence, in conjunction with Amazon Textract. By employing these cloud-based services, the parser could accurately preserve information context by reading in accordance with page formatting, yielding precise and comprehensive results. These results contained a variety of textual, metric, and tabular information found within each document, which was written to a JSON file, and prepared for later uploading to the VS Code program.

The VS Code program hosted the infrastructure required for the prompt engineering portion of the project. Through this environment, the team designed prompts to be passed to the Samantha chatbot, querying the JSON for specific metrics needed for the heuristic. The results of these prompts were then compiled and output as another JSON file, which was uploaded to the PostgreSQL database and converted into a pandas dataframe. The conversion of the JSON to a dataframe allowed the team to extract the necessary metric values to perform the mathematical operations required for the scoring heuristic.

7.3.2 Heuristic Design

To award deal documents an overall score between one and ten, our team first needed to determine which metrics to use as heuristic values. In total, we settled on eleven metrics: %YoY ARR, gross retention, net retention, magic number, gross margin, TAM, %YoY revenue growth, %YoY CARR, TEV/CARR, EBITDA, and CAGR. The amount and variety of metrics chosen was particularly important due to the variety of deal documents slated for analysis. Since investment profiles varied by industry and company model type (e.g., SaaS), our heuristic needed to be broad enough to include metrics that would appear across different types of business financials. Moreover, when evaluating each metric, our team decided to conduct background research on relative performance benchmarks. In doing so, we were able to design customized scoring heuristics for each metric, which were conducive to increasing the accuracy of the eventual overall deal score. Had we scored each metric on the exact same benchmarks, the output would have been inaccurate, as there exist many inconsistencies in data representation within each deal document (e.g., percentages, dollars, fractions, etc.).

To analyze values presented in either decimal format for percentages or in the millions, we implemented a two-step approach. For percentages, we multiplied by one hundred to obtain whole numbers, while for large values in the millions, we divided by one million. This adjustment facilitated ease in code evaluation, sparing us from having to express metric ranges in millions. Additionally, in scoring each individual metric, we incorporated a check for zero values. If a value equated to zero, the score would reflect zero as well, ensuring that it doesn't contribute to the final calculation. This method streamlined the evaluation process and enhanced the overall flow of our analysis.

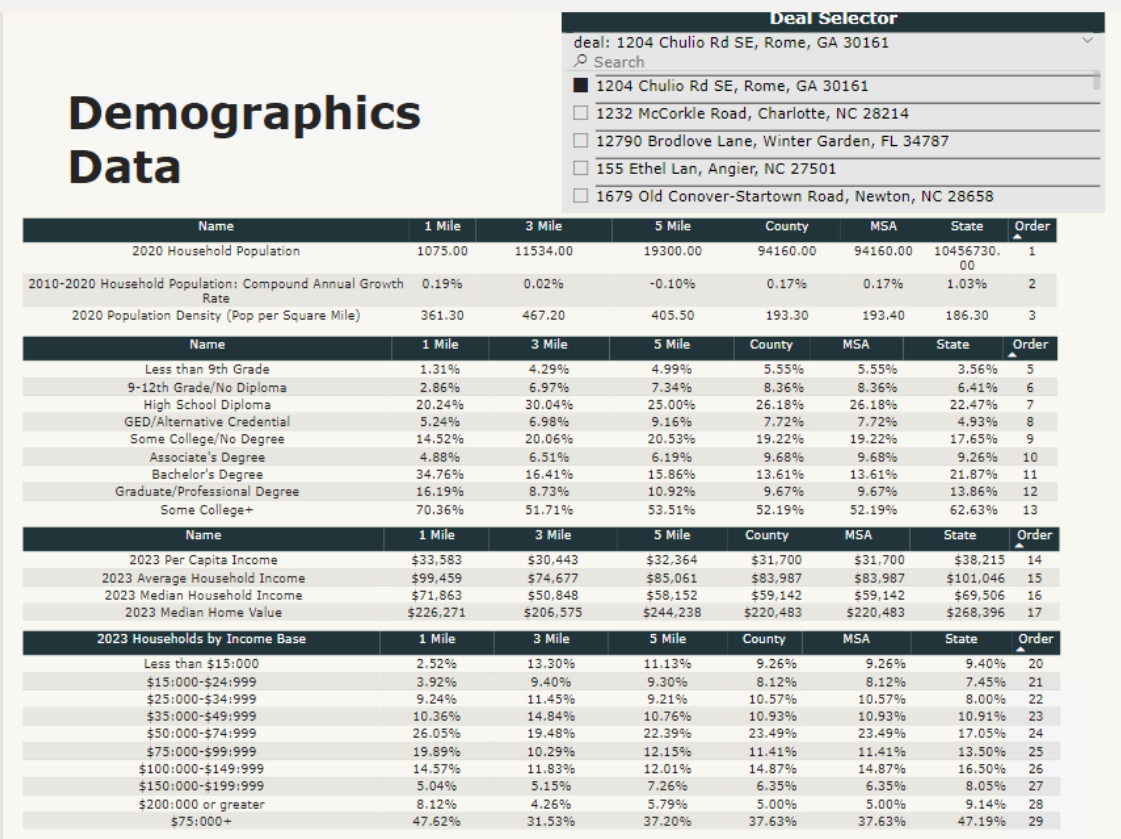
After all individual scores were determined, they were then summed up and divided by the number of values that did not return a score of zero. This ensured that any value not extracted

from the document did not affect the overall final score. The final output was then generated, including the value for each metric, the score each metric received, and the final score, which was presented at the bottom.

8. Results

8.1 Mortgage-Backed Security Desk Dashboard

By leveraging demographic and locational data from ESRI, Redfin market, Zillow, New Home Source, and Multiple Listing Enterprise Solutions, we created a comprehensive Power BI dashboard for an analyst on the residential mortgage-backed security desk. This involved creating multiple dashboard pages with user-friendly features, such as drop-downs to filter the graphs by a particular deal. Additionally, we wrote methods to extract necessary data from external sources such as Zillow and Redfin and correlate them with the company's deal data. These methods generated a new table hosted in Databricks, which provides functionality for scheduling updates in the future. These new tables were then linked to Power BI, where we built out the rest of our graphs. The final dashboard will empower analysts to understand historical trends and make informed decisions on mortgage-backed security deals.



Name	1 Mile	3 Mile	5 Mile	County	MSA	State	Order
2020 Household Population	1075.00	11534.00	19300.00	94160.00	94160.00	10456730.00	1
2010-2020 Household Population: Compound Annual Growth Rate	0.19%	0.02%	-0.10%	0.17%	0.17%	1.03%	2
2020 Population Density (Pop per Square Mile)	361.30	467.20	405.50	193.30	193.40	186.30	3

Name	1 Mile	3 Mile	5 Mile	County	MSA	State	Order
Less than 9th Grade	1.31%	4.29%	4.99%	5.55%	5.55%	3.56%	5
9-12th Grade/No Diploma	2.86%	6.97%	7.34%	8.36%	8.36%	6.41%	6
High School Diploma	20.24%	30.04%	25.00%	26.18%	26.18%	22.47%	7
GED/Alternative Credential	5.24%	6.98%	9.16%	7.72%	7.72%	4.93%	8
Some College/No Degree	14.52%	20.06%	20.53%	19.22%	19.22%	17.65%	9
Associate's Degree	4.88%	6.51%	6.19%	9.68%	9.68%	9.26%	10
Bachelor's Degree	34.76%	16.41%	15.86%	13.61%	13.61%	21.87%	11
Graduate/Professional Degree	16.19%	8.73%	10.92%	9.67%	9.67%	13.86%	12
Some College+	70.36%	51.71%	53.51%	52.19%	52.19%	62.63%	13

Name	1 Mile	3 Mile	5 Mile	County	MSA	State	Order
2023 Per Capita Income	$33,583	$30,443	$32,364	$31,700	$31,700	$38,215	14
2023 Average Household Income	$99,459	$74,677	$85,061	$83,987	$83,987	$101,046	15
2023 Median Household Income	$71,863	$50,848	$58,152	$59,142	$59,142	$69,506	16
2023 Median Home Value	$226,271	$206,575	$244,238	$220,483	$220,483	$268,396	17

2023 Households by Income Base	1 Mile	3 Mile	5 Mile	County	MSA	State	Order
Less than $15:000	2.52%	13.30%	11.13%	9.26%	9.26%	9.40%	20
$15:000-$24:999	3.92%	9.40%	9.30%	8.12%	8.12%	7.45%	21
$25:000-$34:999	9.24%	11.45%	9.21%	10.57%	10.57%	8.00%	22
$35:000-$49:999	10.36%	14.84%	10.76%	10.93%	10.93%	10.91%	23
$50:000-$74:999	26.05%	19.48%	22.39%	23.49%	23.49%	17.05%	24
$75:000-$99:999	19.89%	10.29%	12.15%	11.41%	11.41%	13.50%	25
$100:000-$149:999	14.57%	11.83%	12.01%	14.87%	14.87%	16.50%	26
$150:000-$199:999	5.04%	5.15%	7.26%	6.35%	6.35%	8.05%	27
$200:000 or greater	8.12%	4.26%	5.79%	5.00%	5.00%	9.14%	28
$75:000+	47.62%	31.53%	37.20%	37.63%	37.63%	47.19%	29

Figure 8.1.1: Subset of the ESRI Demographics Data Dashboard

The ESRI Demographic dataset has seven columns of data listing the metric being scored and the distance from the deal that the metric scored at, starting at a 1-mile radius around the deal all the way to the state level. There are a variety of metrics including household population, education, household income, apartment units in the area, time to commute to work, crime index, employment rate, and credit card debt. Each of the tables on this dashboard page, some of which are shown in Figure 8.1.2, go further into detail to allow RMBS investors to better grasp the area surrounding their deals to help determine if the deal is a worthwhile investment.

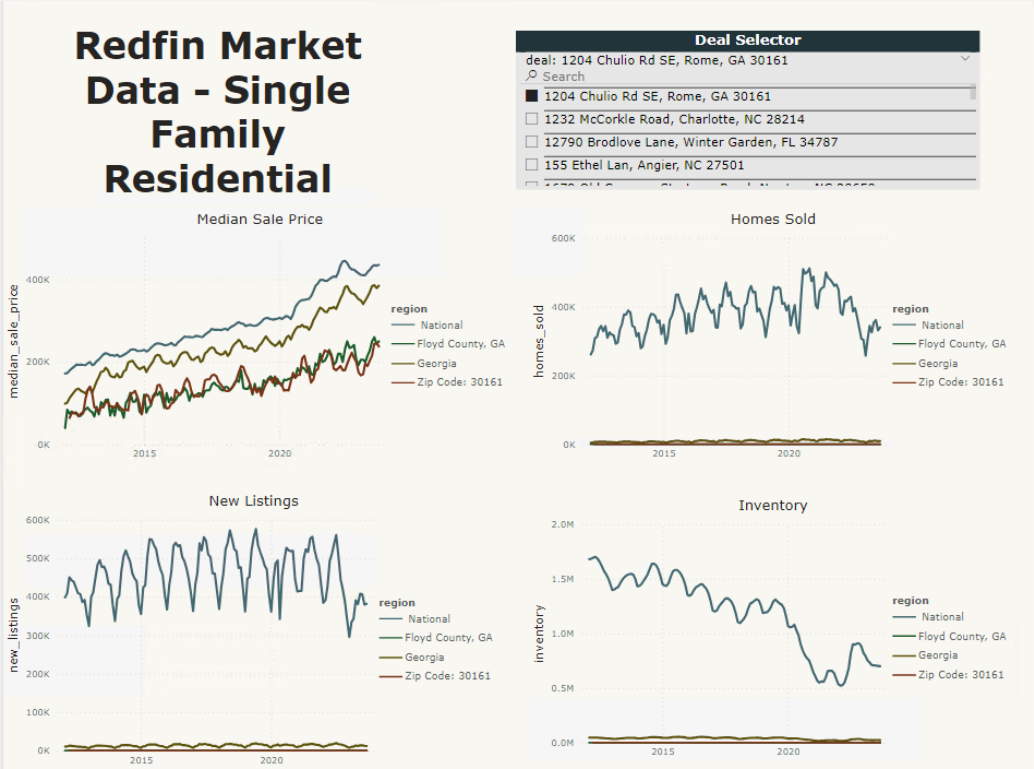


Figure 8.1.2: Subset of the Redfin Market Data – Single Family Residential Dashboard

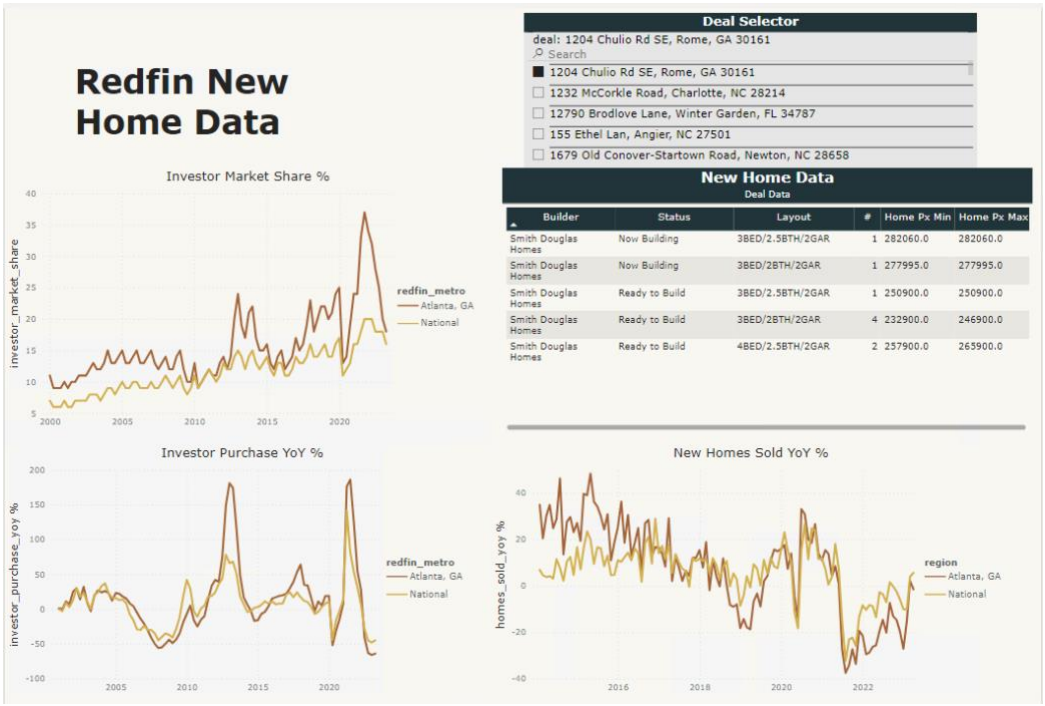


Figure 8.1.3: Subset of the Redfin New Home and New Home Source Dashboard

Redfin is a real estate company that provides investors with mortgage services including data on both the housing market and new homes. The Redfin dataset for single-family residents covers four regions: national, county, state, and zip code. From this data, we developed graphs for the median sale price of a home, number of homes sold, number of new listings, and available properties, and graphs to show the year-over-year percentage for each category. Utilizing other data from Redfin, we designed the New Home dashboard with visuals for market share percentage, investor purchase percentage, new homes sold, new listings, inventory, median sale price, and more. This enables investors to identify favorable deals by comparing housing prices across regions and allows them to make estimates for properties that will or have recently entered the market.

There is also a table created from New Home Source data in Figure 8.1.3 that shows new homes that are being built. It includes the company who is building it, the progress of the building, the planned layout of the home, and the minimum and maximum price the home can sell for.

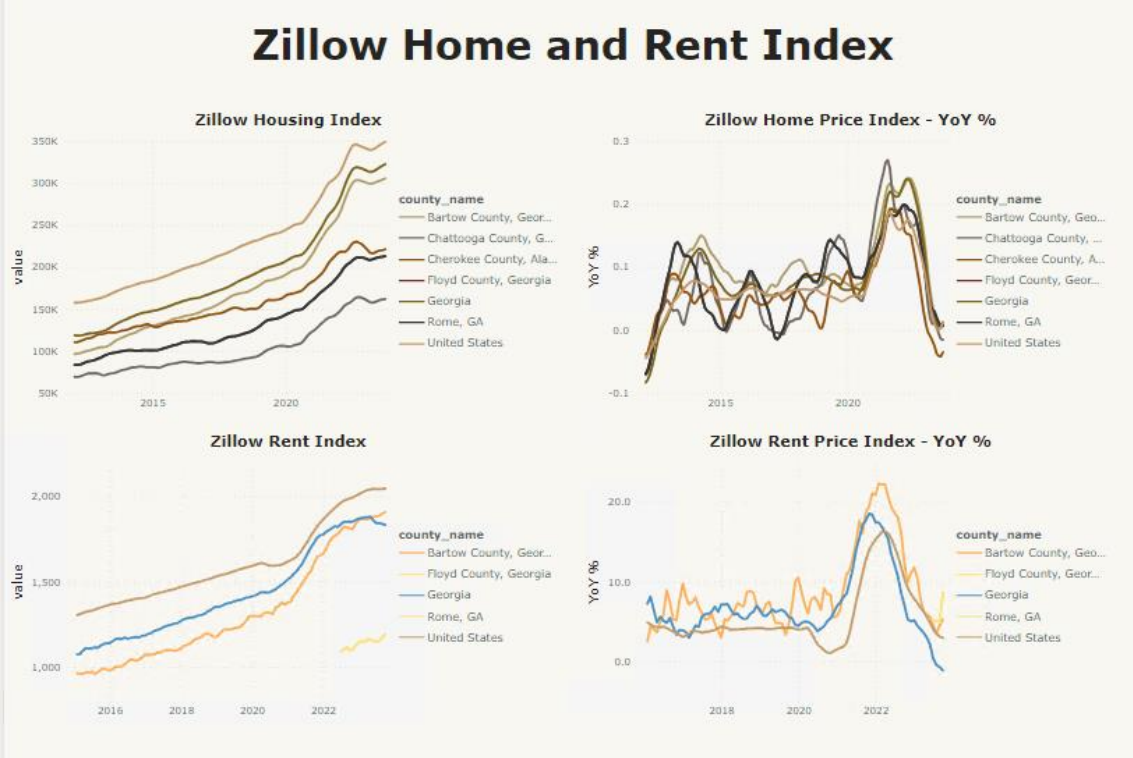


Figure 8.1.4: Zillow Home and Rent Index

The Zillow dashboard provides graphs for the Zillow housing and rent index along with the year-over-year percentage for both as seen in Figure 8.1.4. These specific indices provide insights into the housing market by tracking changes in home values and rental prices over time. The housing index offers an estimate for a typical home in a given area by considering multiple factors, and the rent index measures the median rent in the area and tracks any changes over time. The dashboard also has tables that show the home and rent prices between the years 2019-2023 along with the year-over-year percentage for the pricing change. This information allows investors to understand a customer's point of view when buying a house and can help better inform them whether to invest or not.

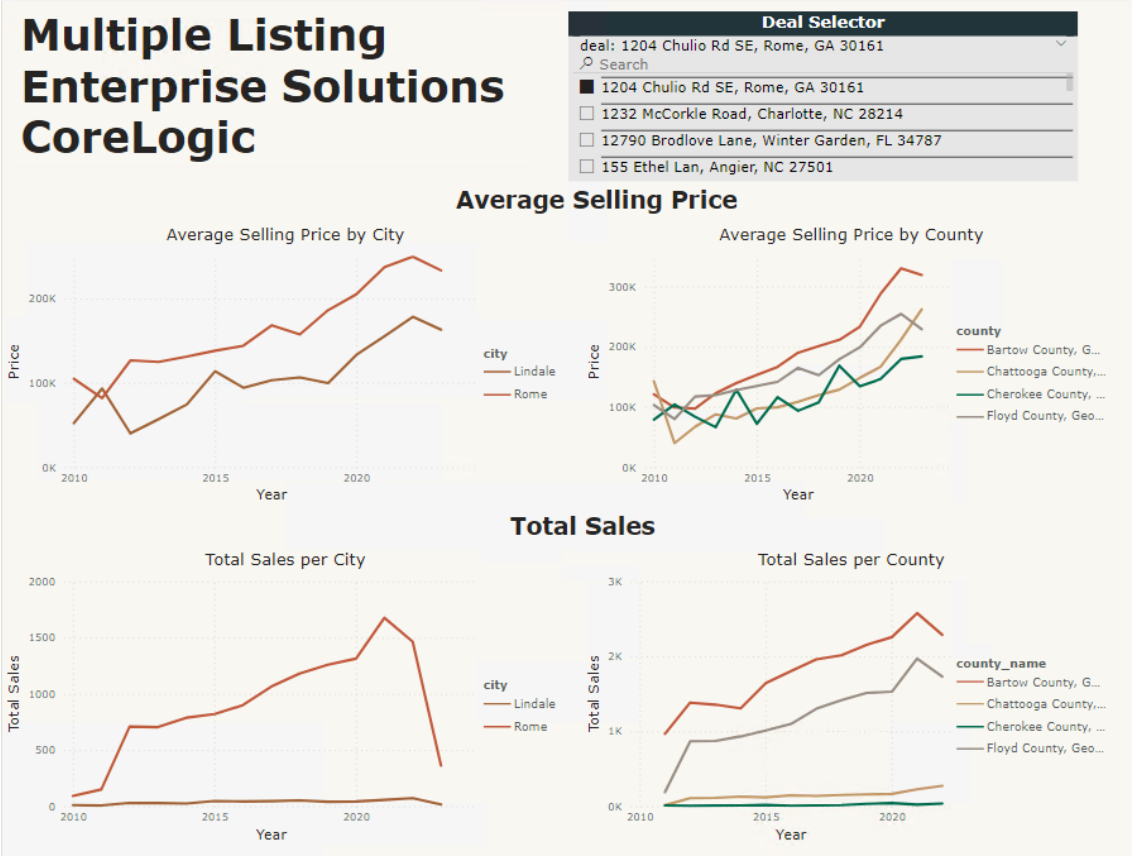


Figure 8.1.5: Subset of the MLS CoreLogic Dashboard

Multiple Listing Enterprise Solutions, or MLS, is a database used by real estate brokers to share information about properties for sale or rent. It allows brokers to see each other's listings of properties and facilitates cooperation and compensation among agents. Utilizing this data, we developed plots that analyze a particular metric, like total sales, depending on the city and county market. All this data allows investors to compare prices and sales with similar deals in the area, thus allowing the investor to understand a typical price for listed homes.

8.2 LLM Deal Appraisal Model

We were tasked with streamlining the review process for potential company investments. Various projects are proposed to the parent company's Investment Committee (IC), each with

extensive documents detailing the goal, company, market, and financial information. The IC screens many deals a day, necessitating a method to quickly identify worthwhile investments to improve screening efficiency.

By utilizing the company's proprietary ChatGPT interface and deal documents, alongside Azure's Form Recognizer, AI Document Intelligence, and Amazon Textract, our team created a model that enables investors to query and generate scores for loaded deal documents based on financial metrics found in the documents. This streamlines the process, allowing the company to reduce the time spent on manual reviews and focus on documents rated highest by the scoring model.

The interface features a user-friendly UI and was developed based on code provided by the company, which we modified to build a specific model. All tests for this model were constructed and executed from Docker, providing local control over the website. After implementing and evaluating the scoring heuristic for each deal, our team settled on the following values to determine the overall score for deals: % YoY ARR Growth, Gross Retention, Net Retention, Magic Number, Gross Margin, TAM, % YoY Revenue Growth, % YoY CARR Growth, TEV/CARR, EBITDA, and CAGR.

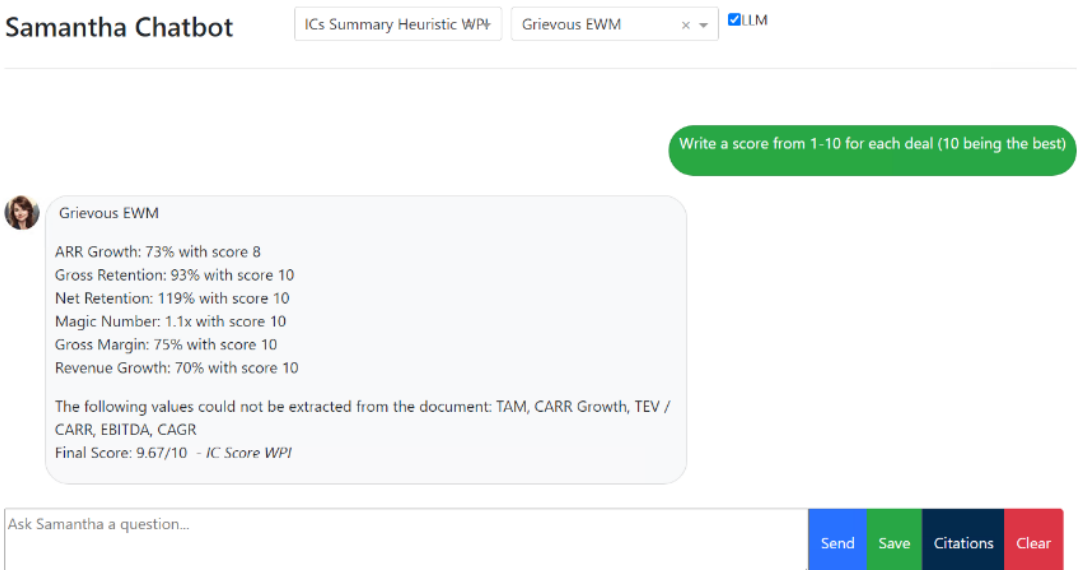


Figure 8.2.1: Internal LLM Deal Appraisal

Users can interact with the model by selecting the 'ICs Summary Heuristic WPI' option from the drop-down menu in the top navigation bar. Additionally, users can choose to specify which document is to be evaluated in the rightmost drop-down menu (if not specified, the model will evaluate all available deals). After configuring these two menus, users can enter their query in the textbox at the bottom of the page; however, it's important to note that their query essentially serves as a starting prompt to run the model and is not considered in the output.

Since the actual scoring prompts run in the backend of the code, the user will always receive a score for the deal. Once the query is processed, the result will be displayed in the format shown above in Figure 8.2.1. This format provides the value extracted from the deal document, the individual score for each metric, and the final overall score for the deal.

To develop the score, we employed prompt engineering techniques to instruct Samantha to extract the correct metrics from the document, format the values appropriately, and then output them as a JSON object. Figure 8.2.2 displays the prompt that we developed. The first section of the prompt instructs the LLM on which metrics to extract from the document and how

to return them in the end. The second section specifies the naming convention for the field names. This is crucial for consistency, enabling us to extract and write values to the Postgres database with uniformity, aligning with the column names. The third section provides instructions to Samantha on all the formatting requirements, such as writing 1.3 million as '1300000'. This ensures that the values we receive will always be in the same format, enabling us to perform mathematical operations on them accurately. Finally, the last section of the prompt explains what to do if there are multiple values for a field and re-specifies the return type.

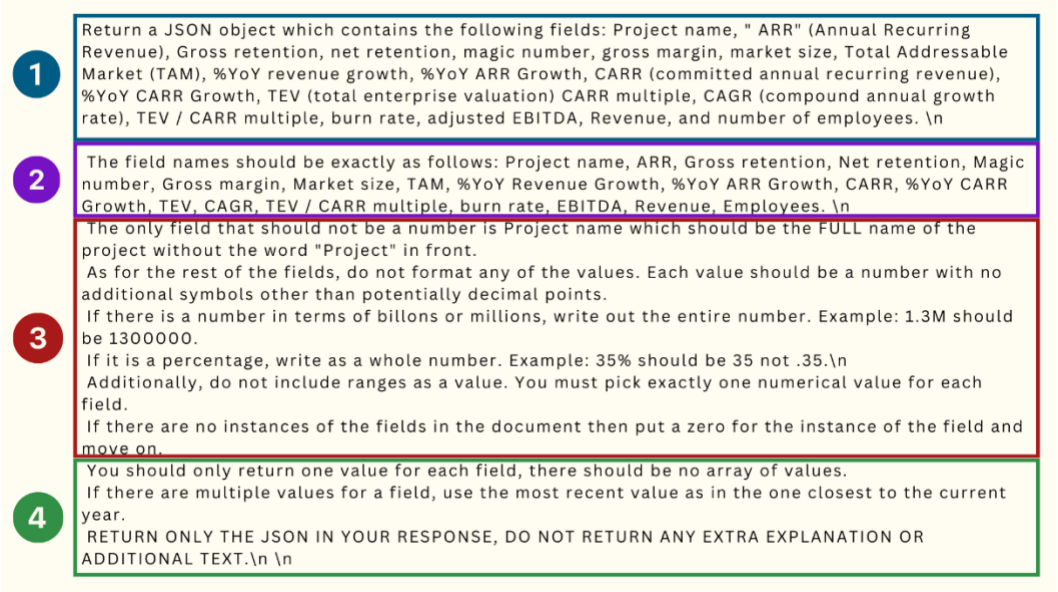


Figure 8.2.2: Prompt to Generate a JSON Object with Desired Metrics

Upon executing the prompt, the LLM extracts the required fields from the selected deal document. Subsequently, we developed a method to read in the JSON files, extracting each metric from its respective field. Our code then assigns a score between 1 and 10 for each metric, utilizing the pre-defined ranges, and constructs a string containing the metric name along with the appended scores. The ultimate score is computed as the average of the extracted metrics, with

the possibility of future adjustments to a weighted average if needed. Finally, the resulting string is output in Samantha's frontend.

In addition to the front-end value and score output, we also wrote the extracted values and scores to a table in Postgres, along with their corresponding deal names. This way, it stores previous runs for later analysis to check for consistency, as LLMs are non-deterministic and may not pull the same value every time.

We ran our final metric ranges and prompt for each of the six deals ten times. The resulting summary statistics are in the table below. The greatest difference in scores was a range of 0.86 points, and the smallest was 0. Some reasons for these discrepancies were metrics being pulled from the document inconsistently and values read incorrectly. For instance, in one of the runs of Clonetrooper IRC, a value of 18% was pulled into '% ARR Growth' even though that metric does not exist in the document; the score for ARR Growth was then given a 2, resulting in a lower overall score compared to the other runs. With some inconsistencies in the scores, we look to continue refining the prompt to account for any values being pulled improperly.

Project Name	Range (min-max)	Mean	Median	Mode
Clonetrooper EWM	7.43 - 8.29	7.847	7.75	7.75
Clonetrooper IRC	8.0 - 8.86	8.64	8.67	8.67
Grievous	9.67 - 9.71	9.68	9.67	9.67
Legitimate EWM	9.0 - 9.11	9.68	9.0	9.0
Palpatine IRC	10.0 - 10.0	10.0	10.0	10.0
Star Destroyer	8.0 - 8.8	8.49	8.8	8.65

Table 8.2: Summary Statistics for LLM Score Precision

The deal with the best score was Project Palpatine (Median: 10), followed by Grievous (Median: 9.67), Legitimate (9.11), Clonetrooper IRC (8.67), Star Destroyer (8.65), and lastly Clonetrooper EWM (7.75). Based on this data, we would suggest filtering out deals like

Clonetrooper EWM because they are less promising. With more data, we could determine a good threshold to exclude deals that should not move forward to the IC's review process. Furthermore, it's important to consider the scoring in relation to the specific market of each company, as score ranges can vary across different markets. For instance, an ARR Growth of 40% may be good for retail but bad for tech companies. Therefore, one should compare scores between deals in the same market and establish thresholds for each category.

Overall, the team feels that the score output provides valuable information to the company. The scores are regular (within a value of 1) and consider key financial metrics for determining company value.

9. Assessment

Dante Amicarella:

Coming into the MQP I had set out three goals for myself to grow in: improve my programming skills, enhance my ability to communicate with professional teams, and understand what working in a fast-paced technical environment looks like. Throughout these projects, I have had to be flexible and play multiple roles within the project team to respond to the changing demands of our sponsor.

Starting from PQP, I was able to insert myself as a developer within my Scrum team and took on the Agile development process in stride. In the first project, I was introduced to Databricks and its functionalities by the company liaison. This environment proved critical as it was the main foundation to create an efficient ETL process for the RMBS data in Power BI, a process which I aided in. These skills allowed me to assume responsibilities for various aspects like Power BI, data extraction, and report writing team in the first project. During the second project, I applied my previous knowledge of working with Docker containers in addition to assisting in the transformation of document data into a JSON file. Financially, I had the opportunity to research financial indicators within deal documents and understand how to extract the information to be used in our model. Within the report, I took on the role of a lead developer, spearheading various chapters. Additionally, I facilitated the revision of my teammates' work to ensure the final product is cohesive and precise. Throughout the process I have learned new software and technical processes in addition to growing my programming skills through my assistance in the parsing of the PDF deal documents.

Alternatively, as a team we took the time to receive criticism from one another to see how we can improve collaboration within the group. These refined soft skills have empowered me to facilitate more impactful and productive conversations within the group and allowed me to become a better teammate.

Maya Liao:

Entering this project, my objectives encompassed the growth of technical skills (coding and data analysis), improved client communication, and enhanced project time management. Upon completing the project, I found that I had achieved these goals to varying extents.

In terms of technical skills, my involvement in the project included working on backend data pipelines and utilizing Power BI for the RMBS Dashboard. While I had prior experience coding data pipelines in R and Python, this project introduced me to the Databricks and PostgreSQL environments. Spearheading the establishment of workflows for the RMBS dashboard not only acquainted me with the Databricks user interface but also exposed me to a range of data cleaning and transformation operations necessary for producing accurate datasets. Additionally, I played a smaller yet valuable role in working with Power BI, gaining insights into its functionalities, particularly since my prior experience had been with Tableau. Similarly, in the context of the LLM Deal Appraisal Prompt segment, I encountered a similar scenario working on backend data pipelines and the app interface. Leveraging my familiarity with Databricks and VS Code from the RMBS Dashboard and other previous experiences, I adeptly handled the required tasks in this domain.

Alternatively, assessing success in terms of client communication and project time management goals presented more challenges. Throughout the project, I served as the Product Owner, entailing responsibilities such as communicating with key stakeholders (the company

liaison) and overseeing the project's progress. Unlike a typical AGILE Scrum team, I also took on the role of a Developer, requiring contributions to the design of deliverables. Consequently, when communicating with the sponsor, most of my time was dedicated to clarifying project objectives and addressing any deliverable blockers the team encountered. Following discussions within our team, we concluded that this approach worked effectively, highlighting my ability to lead meetings and identify project requirements. Regarding project time management, it posed considerable difficulties due to the condensed timeline for our project. The delayed onboarding and an earlier final deliverable due date resulted in significantly shortened durations for our development and testing stages. The challenges escalated during the LLM Deal Appraisal portion of the project, encountering significant development blockers over the Thanksgiving weekend. The combination of the holiday vacation and the pause in deliverable development, necessitated by the need for further guidance from our sponsor, resulted in the loss of three days' worth of technical work. Looking back at the situation, I believe I handled most of the project's time management well but could have avoided the Thanksgiving debacle if I had been more attentive to the developmental work being done for that portion of the project. In the future, I will make sure to plan further ahead to avoid developmental blockers that coincide with working hour conflicts.

Nathan Shemesh:

I set out three goals for myself when coming into this MQP and I wanted to improve my technical skills, effectively communicating with both the sponsor and the project team and make improvements in my time management. Throughout the project I was able to achieve these goals. During the first part of the project where the team worked on making the RMBS dashboard I had the opportunity to work on linking the database from DBeaver to Power BI. It was my first time

working with Power BI where I learned how to make tables and manipulate the formatting of the data using DAX formulas while also learning how to filter the data by the selected deal using the slicer. For the second part of the project where we were utilizing the company's internal LLM to appraise investment deals, I learned a great deal about prompt engineering and how to write effective prompt to a LLM so that it can analyze the documents. During this part of the project, I utilized both VS Code and Docker to run and test these prompts and while I had a great deal of experience with the VS Code IDE it was my first-time using Docker to build and run an application. I also gained more experience with coding in Databricks as one of the challenges that we came across during the project was that when reading in the document from the JSON file it did not format the tables in the document well enough for the LLM to interpret so I had to extract the table data into a dataframe and load it into the index that we were using so that the LLM could interpret it.

During the seven weeks of this project my team members and I have had daily meetings with our liaison in the company where we would communicate our progress with the project as well as any potential problems we had. Furthermore, our team would meet up on a regular basis to work on the project and by having all of us working in the same space it facilitated better communication among team member and we were able to solve problem more effectively that if we were to work remotely from our homes. By utilizing the daily stand up it gave me the chance to speak up about my progress during the project and any challenges I had, and it was nice to have a dedicated time to do this as the stand-up was well organized and it allowed for my workflow to not be impeded.

Since the project was only seven weeks, time management played a big part in our teams' success. We often worked from 9 to 5 pm, this along with well-defined deadlines for when things

were due provided me with a clear understanding of how to divide my time so that I could complete the work that I was assigned.

Sarah LaRusso:

During this project, I was able to meet all my goals and gain an invaluable set of technical skills and knowledge. Notably, I became proficient in Power BI, a new tool for me, and applied recently acquired SQL skills to a company project. I also broadened my financial knowledge. I explored actual deal documents gaining insights into the intricacies of company investments and got the opportunity to speak with Professor Dunbar to receive guidance on understanding and selecting business metrics. In addition to learning about financial data, I expanded my knowledge on LLMs and prompt engineering, an area I am quite interested in. My mindset on GPT models certainly changed as well; before the project I was resistant to using ChatGPT, but now I cannot imagine not using it to answer questions and work through debugging.

In addition to honing technical skills, I also enhanced my soft skills. Collaborating with a group of three colleagues in a fast-paced environment doing 9-5 workdays was a drastic change from the usual three classes WPI students have at once. I had to adjust to working on one task for an extended period and interacting with the same partners repeatedly. This structure demanded effective communication, honesty, and agility to foster a productive environment. This experience not only improved my collaborative skills but also emphasized the importance of transparent communication in achieving shared objectives. Furthermore, it provided me the opportunity to assume responsibility overseeing our Agile practices, allowing me to apply and develop my leadership skills and effective communication strategies.

Team:

Entering the project, our team were unfamiliar with one another. However, throughout the 14 weeks, we developed into a functionally dynamic group, providing continuous support for each other. During the project, our team diligently adhered to our Agile practices, ensuring we remained aligned with our roles and sprint goals. Through the effective implementation of daily stand-up, sprint planning, sprint review, and sprint retrospective meetings, the entire team maintained a clear understanding of the expectations we set on a week-to-week basis. This approach facilitated a seamless workflow and enhanced collaboration among team members. The dynamic nature of our workspace fostered an environment where each member could authentically express themselves in a professional setting, leveraging their unique skills and assets brought to the team.

At the beginning of the semester, our goal was to create a functional deliverable that aids the alternative investment firm's investment strategies and resources. Initially, we anticipated focusing primarily on a project related to the alternative investment firm's mortgage-backed security data. However, upon our introduction to the liaison, it became evident that our expectations were incorrect. With this realization, our initial project served as a launchpad for our main LLM Deal Appraisal project. As a team, we were adaptable and able to change gears quickly to understand the new task at hand. We collaborated consistently to comprehend various aspects, including utilizing Databricks, prompting the company's large language model, interpreting financial data, templating Power BI, and most importantly, how to work with one another. While our goals fluctuated on a week-to-week basis, our main overarching deliverable was always at the forefront of our work. In the end, we demonstrated resiliency by working on revolutionizing technology to aid our sponsor.

10. Business and Risk Management

10.1 Overview

The alternative investment company faces pivotal moments requiring risk-based decisions that may result in either high or low rewards depending on the outcome. Regarding our project, the company took a calculated risk by engaging with us. The potential success of our project promises rewards far surpassing the associated risks. These strategic decisions not only propel the alternative investment company forward but also cultivate an environment allowing for measured freedoms within the organization.

10.2 Risk Culture

Throughout our project, our team closely observed the risk culture embedded in the daily operations of the alternative investment company. During the onboarding process, we were required to submit both our United States passport and Social Security card for a comprehensive background check on all team members. This stringent security protocol underscored the company's dedication to prioritizing operational safety, particularly when bringing on potentially risky employees. Following the onboarding phase, our team maintained consistent collaboration with the security desk to obtain clearances for all applications essential to project tasks. These ongoing measures aimed to minimize the risk of employees downloading potentially harmful software, effectively safeguarding the integrity of the company's data.

Moreover, when we accessed the project data through Azure Web Services, our company liaison promptly granted us access, recognizing the risk-averse nature of our initial project. These swift actions enabled the team to seamlessly explore the data and advance toward the product. In a parallel vein, the company also hosted a localized instance of ChatGPT named

Samantha. The active encouragement of utilizing this chatbot underscored the company's forward-thinking stance and acceptance of utilizing artificial intelligence within the workplace.

Despite this acceptance of innovative technology, our team, particularly in the initial project phase, consistently prioritized validating imported data into Power BI. This diligence was key to ensuring that any company analyst using the dashboard could confidently interact with and derive insights from high-quality, reliable data. This commitment to data accuracy aligned with the company's dedication to ensuring all data used for decision-making was risk-free in its analytical processes.

In our second project, the company displayed a commendable level of risk tolerance by fostering an environment that encouraged us to consult with professors across WPI. This collaborative approach allowed us to seek valuable feedback on our financial models within our heuristic, ensuring the robustness of our analyses. Given the innovative nature of our second project, the company liaison also provided continuous assistance during workdays. This latitude enabled us to make mistakes in refining the heuristic, acknowledging the potential of our efforts to yield a model capable of saving the Investment Committee numerous hours weekly. While navigating the delicate balance between financial risks and efficiency, this risk-tolerant approach proved instrumental in our pursuit of a robust and time-saving model.

Overall, the scope of our entire Major Qualifying Project weighed heavily on improving both the parent and the alternative investment firms processes within their workspace. Given that both projects were designed to aid the work of employees in both firms, our work was risk-averse, as they did not require the halting or removal of pre-existing procedures.

10.3 Other Risks

Participating in financial risks within a financial institution carried the inherent potential to harm its reputation if investments turned unfavorable. Despite the significance of this reputational risk, our project encountered an additional challenge: the unavailability of actual parent company metrics crucial for developing our heuristic. While this risk-averse approach did not allow us to customize our model directly for the company, it did empower the group to craft a model that became risk-free once adapted to the company's criteria.

11. Future Work

Given the chance to extend our collaboration with the alternative investment company, our team would ambitiously take on both projects, introducing fresh elements into each.

In our initial project, the RMBS Dashboard, the focus was on the accuracy of the data, leaving room for enhancements within the user interface (UI). Given the opportunity to extend our partnership, a key area we would address is the design of the user interface, introducing a cleaner and more polished template to elevate the overall dashboard aesthetics. To accomplish this, our first step would be to add an introduction page that would articulate the purpose of RMBS Dashboard along with a comprehensive overview of the data included. This introductory page would set the stage for analysts and offer clear insights into the dashboard's objective. Additionally, we envision implementing functional headers for each figure title, ensuring that updating filters seamlessly cascades changes across all figure titles. This approach would not only enhance the visual appeal but also reassure the user that they are, in fact, looking at the correct data across all the tables and graphs. A finishing touch would involve incorporating the company's official logo on every page of the dashboard, reinforcing its identity and ownership over the data within.

In addition to enhancing the UI of the dashboards, we would also undertake the tasks of automatic updates and developing more graphs. Initially, we would establish a schedule through Databricks to automatically update the tables and connected dashboard every quarter with new data. This approach ensures that the plots the user sees are up-to-date and accurate. Subsequently, we would explore other datasets that the firm has access to and identify additional

graphics that we consider useful to an RMBS worker. We would then proceed to build out those graphics to help inform their decisions.

As for the LLM Deal Appraisal Prompt Project, our team was unable to incorporate the parent company's investment heuristic due to security protocols. In the future, if the team is cleared to utilize the heuristic, we will incorporate the new metrics across all Investment Committee deal sheets to enhance the appraisal accuracy of the large language model. Additionally, with access to all possible resources, we would compile a more extensive list of IC deal documents. This would allow us to test our model on a broader set of data and continue refining it. The addition of more documents will also provide us with enough information to determine the threshold of which deals should move forward to the IC for review. Furthermore, we would take on the reading of other deal documents such as CIS, QLDP, and QLD documents. Through reading these other three types of documents we would gather an understanding of which metrics are important as these documents would require specified designed prompts for each type of document to extract as much valuable financial information. Then we would follow a similar methodology as done for the IC documents to complete our scoring heuristic. These insights would allow the LLM to centralize on the correct valuation for each document in accordance with the metrics given.

12. Conclusion

In the realm of finance, particularly in real estate, investment management is crucial for strategically allocating assets to maximize returns and mitigate risks for individuals, institutions, or funds. Collaborating with an alternative investment company, our team had a twofold project approach: to create an automated process for viewing real estate investment data and to utilize artificial intelligence to streamline the appraisal of deal documents, saving company analysts time in their functional work with potential and historical investment deal data. By leveraging Agile practices, we successfully conducted weekly sprints to bring the project to completion.

In the RMBS Dashboard project, our team crafted a comprehensive Power BI dashboard that illustrated historical deal data, unraveling the complexities of mortgage-backed securities across different states, cities, and zip codes. This project was accomplished using Azure Data Lake, Databricks, PostgreSQL, DBeaver, Python, pandas, PySpark, and Power BI. Moreover, we created an effective extract, transform, and load (ETL) pipeline for data to generate twenty-five distinct data tables in Data Lake, allowing for scheduled updates. The resulting dashboards encompassed robust data from sources like Redfin, Zillow, and ESRI demographics, offering flexible analysis of mortgage-backed securities across the United States.

The acquisition of the alternative investment firm by a new parent company led us to our second project, the LLM Deal Appraisal Prompt, involving the use of artificial intelligence to appraise deal documents. Through a blend of background research on interpreting financial information and meticulous data parsing, the team designed a heuristic to assign quantitative values to documents using a large language model. This novel project was accomplished using the following tools: Databricks, Amazon Textract, Azure Form Recognizer, Python, pandas, VS Code, LangChain, Azure Cognitive Search, and ChatGPT-4.

In summary, our team is confident that the advanced technology-driven deliverables we've created will greatly enhance operations for both the parent company and the alternative investment firm. Our RMBS dashboards and LLM Deal Appraisal Prompt not only empower comprehensive analysis of mortgage-backed securities but also streamline the company's investment review process. These deliverables are poised to strengthen both firms' approaches to navigating the financial landscape.

References

ARR Growth Rate. (2023, October 12). https://www.metrichq.org/saas/arr-growth-rate/

Ascent. (2022). *Calculating Year-over-Year Growth*. https://ascent.sba.gov/d9/c3/2afa68a841189d27ef5d9c374d85/your-business-financial-strategy-4-2-financial-kpis-tool.pdf

Atlassian. (2019, January 4). *Product Backlog Explained [+ Examples]*. Atlassian. https://www.atlassian.com/agile/scrum/backlogs

Atlassian. (2023a). *Jira | Issue & Project Tracking Software*. Atlassian. https://www.atlassian.com/software/jira

Atlassian. (2023b). *Trello Guides: Help Getting Started With Trello | Trello*. https://trello.com/guide

AWS. (2023). *What is Spark? - Introduction to Apache Spark and Analytics - AWS*. Amazon Web Services, Inc. https://aws.amazon.com/what-is/apache-spark/

Bloomenthal, A. (2022, July 9). *Gross Margin: Definition, Example, Formula, and How to Calculate*. Investopedia. https://www.investopedia.com/terms/g/grossmargin.asp

Boykin, R. (2023, August 31). *The Great Recession's Impact on the Housing Market*. Investopedia. https://www.investopedia.com/investing/great-recessions-impact-housing-market/

Boyte-White, C. (2023, May 24). *Revenue vs. Income: What's the Difference?* Investopedia. https://www.investopedia.com/ask/answers/122214/what-difference-between-revenue-and-income.asp

Bureau of Economic Analysis. (2021, January 29). *U.S. Economy at a Glance | U.S. Bureau of Economic Analysis (BEA)*. U.S. Economy at a Glance. https://www.bea.gov/news/glance

CFI, T. (2023). *CAGR*. Corporate Finance Institute. https://corporatefinanceinstitute.com/resources/valuation/what-is-cagr/

Chen, J. (2023, September 29). *Securitization: Definition, Pros & Cons, Example*. Investopedia. https://www.investopedia.com/terms/s/securitization.asp

ClouldBlue. (2023, July 24). *Contracted Annual Recurring Revenue (CARR) | CloudBlue*. https://www.cloudblue.com/glossary/contracted-annual-recurring-revenue-carr/

Coghlan, E., McCorkell, L., & Hinkely, S. (2018, September 19). *What Really Caused the Great Recession?* Institute for Research on Labor and Employment. https://irle.berkeley.edu/publications/irle-policy-brief/what-really-caused-the-great-recession/

Consumer Financial Protection Bureau. (2017, February 24). *What is a subprime mortgage?* Consumer Financial Protection Bureau. https://www.consumerfinance.gov/ask-cfpb/what-is-a-subprime-mortgage-en-110/

Databricks. (2023, November 15). *What is Databricks?* https://docs.databricks.com

datascience904. (2019, August 23). Word Embeddings. *Data Science*. https://datascience904.wordpress.com/2019/08/23/word-embeddings/

DBeaver. (2023, November 19). *DBeaver Community | Free Universal Database Tool*. https://dbeaver.io/

Docker. (2023). *What is a Container? | Docker*. https://www.docker.com/resources/what-container/

Fernando, J. (2023, May 24). *Compound Annual Growth Rate (CAGR) Formula and Calculation*. Investopedia. https://www.investopedia.com/terms/c/cagr.asp

Fidelity. (2023, November 9). *Standard Deviation Indicator—Fidelity*. https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/standard-deviation

Frankel, K., Zopf, J., & Stuart, J. (2023, November 8). *ARR Growth Rate*. https://www.parative.com/glossary-term/arr-growth-rate

Gareth. (2021, May 13). The 5 Scrum Ceremonies Explained for Remote Teams. *Parabol*. https://www.parabol.co/blog/scrum-ceremonies-for-remote-teams/

Ghanizada, S. (2023, April 14). *Market Size: What is TAM, SAM, & SOM? | Carta*. https://carta.com/blog/market-size/

Google. (2023, August 8). *Introduction to Large Language Models | Machine Learning*. Google for Developers. https://developers.google.com/machine-learning/resources/intro-llms

Groccia, S. (2022, February 2). *What are Gross and Net Revenue Retention? | Mosaic*. https://www.mosaic.tech/financial-metrics/net-and-gross-dollar-retention

Hayes, T. (2022, April 24). *What is net revenue retention & how to calculate it*. https://www.paddle.com/blog/net-revenue-retention-the-new-benchmark-metric-for-saas

Introducing ChatGPT. (2022, November 30). Introducing ChatGPT. https://openai.com/blog/chatgpt

Johnson, S. (2023, January 31). *Investment Firms & the State of Home Buying in the US*. BillTrack50. https://www.billtrack50.com/blog/investment-firms-and-home-buying/

Kagan, J. (2023, April 29). *Mortgage-Backed Securities (MBS) Definition: Types of Investment*. Investopedia. https://www.investopedia.com/terms/m/mbs.asp

Karami, F. (2023, July 10). *Decoding the Magic of Self-Attention: A Deep Dive into its Intuition and Mechanisms*. Medium. https://medium.com/@farzad.karami/decoding-the-magic-of-self-attention-a-deep-dive-into-its-intuition-and-mechanisms-394aa98f34c5

Kenton, W. (2022a, June 20). *Adjusted EBITDA: Definition, Formula and How to Calculate*. Investopedia. https://www.investopedia.com/terms/a/adjusted-ebitda.asp

Kenton, W. (2022b, August 23). *Total Enterprise Valuation (TEV): Definition, Calculation, Uses*. Investopedia. https://www.investopedia.com/terms/t/tev.asp

Kenton, W. (2023, April 24). *What are Financial Securities? Examples, Types, Regulation, and Importance*. Investopedia. https://www.investopedia.com/terms/s/security.asp

Klipfolio Inc. (2023). *Understanding the Difference Between ARR and MRR | Klipfolio*. https://www.klipfolio.com/resources/kpi-examples/saas/mrr-vs-arr

Lee, T. B. (2023, November 15). *Large language models, explained with a minimum of math and jargon*. https://www.understandingai.org/p/large-language-models-explained-with

Lido. (2023, December 4). *Gross Revenue Retention (GRR)—Lido.app*. https://www.lido.app/metrics/gross-revenue-retention

Liu, J., Liu, A., Lu, X., Welleck, S., West, P., Le Bras, R., Choi, Y., & Hajishirzi, H. (2022, September 28). *Generated Knowledge Prompting for Commonsense Reasoning*. https://arxiv.org/pdf/2110.08387.pdf

Mansa, J. (2023, May 24). *Compound Annual Growth Rate (CAGR) Formula and Calculation*. Investopedia. https://www.investopedia.com/terms/c/cagr.asp

Marr, B. (2023, May 30). *10 Amazing Real-World Examples Of How Companies Are Using ChatGPT In 2023*. Forbes. https://www.forbes.com/sites/bernardmarr/2023/05/30/10-amazing-real-world-examples-of-how-companies-are-using-chatgpt-in-2023/

Microsoft Azure. (2023). *Cloud Computing Services | Microsoft Azure*. https://azure.microsoft.com/en-us

Microsoft. (2023a, June 13). *Power BI - Data Visualization | Microsoft Power Platform*. https://www.microsoft.com/en-us/power-platform/products/power-bi

Microsoft. (2023b). *Visual Studio Code—Code Editing. Redefined*. https://code.visualstudio.com/

Morgan, A. (2023, August 14). *Explainable AI: Visualizing Attention in Transformers - MLOps Community*. https://www.comet.com/site/blog/explainable-ai-for-transformers/, https://mlops.community/explainable-ai-visualizing-attention-in-transformers/

Mosaic. (2023). *Guide to Calculating the SaaS Magic Number—Mosaic*. https://www.mosaic.tech/financial-metrics/magic-number

Mozilla. (2023, July 3). *Working with JSON - Learn web development | MDN*. https://developer.mozilla.org/en-US/docs/Learn/JavaScript/Objects/JSON

Nvidia. (2023, March 24). *What are Large Language Models? | NVIDIA Glossary*. Large Language Models Explained. https://www.nvidia.com/en-us/glossary/data-science/large-language-models/

O'Dore, J. (n.d.). *What the heck is adjusted EBITDA and why is it so darn important? | CFO.University*. Retrieved November 22, 2023, from https://cfo.university/library/article/what-the-heck-is-adjusted-ebitda-and-why-is-it-so-darn-important

Open AI. (2023, March 27). *GPT-4 Technical Report*. https://arxiv.org/pdf/2303.08774.pdf

Product Plan. (2023, November 16). *Annual Recurring Revenue*. https://www.productplan.com/glossary/annual-recurring-revenue/

Riedel, S., Kiela, D., Lewis, P., & Piktus, A. (2020, September 28). *Retrieval Augmented Generation: Streamlining the creation of intelligent natural language processing models*. https://ai.meta.com/blog/retrieval-augmented-generation-streamlining-the-creation-of-intelligent-natural-language-processing-models/

Saravia, E. (2023, October 2). *Prompt Engineering Guide*. https://www.promptingguide.ai/

Schulhoff, Sander, & Community Contributors. (n.d.). *Learn Prompting: Your Guide to Communicating with AI*. Retrieved November 9, 2023, from https://learnprompting.org/docs/intermediate/generated_knowledge

Schwaber, K., & Sutherland, J. (2020, November). *The Scrum Guide*. The Scrum Guide. https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-US.pdf#zoom=100

Scott, G. (2022, June 20). *Adjusted EBITDA: Definition, Formula and How to Calculate*. Investopedia. https://www.investopedia.com/terms/a/adjusted-ebitda.asp

Shark Finesse Ltd. (2021, October 19). *What is a good CAGR | Shark Finesse Blog | Shark Finesse Blog*. Shark Finesse. https://sharkfinesse.com/blog/what-is-a-good-cagr

Simpson, S. (2022, December 2). *A Career in Real Estate Portfolio Management*. Investopedia. https://www.investopedia.com/articles/financialcareers/09/real-estate-portfolio-management.asp

Stripe. (2023, November 14). *Net revenue retention vs. Gross revenue retention | Stripe*. https://stripe.com/resources/more/net-revenue-retention-vs-gross-revenue-retention

Vanderjack, B. (2015). *The Agile Edge: Managing Projects Effectively Using Agile Scrum*. Business Expert Press. http://ebookcentral.proquest.com/lib/wpi/detail.action?docID=2145193

Verlaque, M. (2023, July 20). *ARR (Annual Recurring Revenue): How to Calculate It | SaaS Academy*. https://www.saasacademy.com/blog/how-to-calculate-annual-recurring-revenue

Wall Street Prep. (2022, January 4). *Total Addressable Market (TAM)*. Wall Street Prep. https://www.wallstreetprep.com/knowledge/total-addressable-market-tam/

Weinberg, J. (2013, November 22). *The Great Recession and Its Aftermath | Federal Reserve History*. The Great Recession and Its Aftermath. https://www.federalreservehistory.org/essays/great-recession-and-its-aftermath#rise

Wing. (2023). *Understanding TAM (and SAM and SOM)*. https://www.wing.vc/docs/product/understanding-tam-and-sam-and-som