Project Number:51-JAG-002-5/<br>51-JAG-004<br>51-JAG-005

## ADVANCED PLACEMENT PROGRAM AND COLLEGE PERFORMANCE

An Interactive Qualifying Project<br>Submitted to the Faculty<br>of the

WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science


Date: May 8, 2000
Approved:

1. AP testing
2. college performance
3. comparison


## Abstract

High School advanced placement testing and courses are explained. Analysis is done to identify who at WPI is an advanced placement (AP) student. Then analysis is done to determine if AP students perform better at WPI than non-AP students. The hypothesis is that the AP students do in fact perform better than the non-AP students.

## Acknowledgements

We would like to thank the following people:

## Mr. David W. Galvin

For his help in obtaining the data we used for this project.

## Professor John Goulet

Our advisor, for his help and advice throughout our project.
INTRODUCTION ..... 5
BACKGROUND ..... 6
Advanced Placement Program ..... 6
GRADE POINT AVERAGE ..... 10
MICROSOFT EXCEL ..... 11
MAIN HYPOTHESIS ..... 13
METHOD: ..... 15
IDENTIFIERS ..... 15
AP MATH ..... 18
AP CHEMISTRY ..... 19
AP COMPUTER SCIENCE ..... 20
GPA I AND GPA II ..... 20
HYPOTHESIS TESTING ..... 22
CONCLUSIONS DRAWN FROM ANALYSIS ..... 24
MATH ..... 25
DIFFERENTIAL EQUATIONS ..... 28
Calculus IV ..... 29
CHEMISTRY ..... 31
COMPUTER SCIENCE ..... 32
IQP AND MQP ..... 33
SAT DATA ..... 34
GENDER ..... 35
SUMMARY: ..... 36
MATH ..... 37
CHEMISTRY ..... 39
COMPUTER SCIENCE ..... 40
CONCLUSION ..... 41

## Introduction

The Advanced Placement (AP) program run by the College Board and administered by many High Schools provides students with the opportunity to take college level courses. The program also provides exams at the end of the course. Many students coming into WPI have taken AP courses and exams in mainly three subject areas. AP students have a small jump on the rest of the population because they are able to start in higher level courses when they arrive at school. Although WPI does not have a specific program for AP students, meaning it is up to the students to decide what class to take, most AP students skip over one or two classes in the given subject.

This report provides WPI with data and analysis that tries to prove the hypothesis that AP students perform better at the college level than non-AP students. We investigated 3 subjects: math, chemistry, and computer science. These 3 subjects are all offered in the AP program. By analyzing students' (AP and otherwise) performances throughout their college careers we will show whether or not the hypothesis is true. The analysis is done and presented in tables and charts.

We have obtained data on 588 students at WPI ranging from 1995 to 2000, some of which have taken the AP courses and some not. This data provides us with the information from one whole class through their for years at WPI. We used 2 different methods to analyze our data, a grade distribution, and an unofficial GPA. From our data we have analyzed AP Calculus, AP Chemistry and AP Computer Science students, experimental group, against students who have not taken the courses, control group. Overall we have provided a better understanding of how Advanced Placement students perform at this school.

## Background

## Advanced Placement Program

The Advanced Placement (AP) Program is a program run by the College Board. It is offered by over 14,000 high schools all over the world. In 1999 over 1 million exams were taken by almost 700,000 students. The AP Program allows high school students to challenge themselves by taking college level classes while in high school. The program offers 32 courses in 18 subject areas. Some of these subjects include calculus, physics, chemistry, biology, history, and computer science. After the course work has been finished the students are then allowed to take AP exams in the respective subject.

The College Board offers these exams at the end of each school year in May. Students may take more than one exam, depending on the number of AP courses taken. The cost for taking each exam is $\$ 76$. The College Board does offer financial aid for those students in the Program that need it in order to take the test. Although the exam does cost money, it could, if the student does well enough on the exam, save them from taking entry level courses in their first year in college and thus saving them much more money in the long run. The students could possibly graduate earlier, given they skip these entry level courses. This is a great money saver, because the students don't have to pay for tuition for the semesters they are missing, while the students who didn't skip these classes are paying for them.

We will be concerned with three subjects: Calculus, Chemistry, and Computer Science. For some of these subjects there are different courses offered. For example, in calculus there are two courses offered: Calculus AB and Calculus BC. Similarly, the College Board offers Computer Science A and Computer Science AB. Each of these courses has a separate exam, each of which focus on different aspects of the subject. This will be explained in more detail so as to get a better understanding of what the AP courses and AP exam are.

As mentioned before there are two separate courses offered in Calculus. In order for a student to participate in the AP calculus course there are certain prerequisites that must be met. Before taking the course a student must have sufficient knowledge in Algebra, Geometry, Trigonometry, Analytic Geometry, and Elementary Functions. Calculus AB can be offered by any school with a curriculum that includes all of the prerequisites for a year's course in calculus. The material included in the Calculus AB exam is to be taught over a full high school academic year. It is possible to spend some time on elementary functions and still cover the Calculus AB curriculum within a year. However, if students are to be adequately prepared for the Calculus AB examination, most of the year must be devoted to topics in differential and integral calculus (calculus I and II at WPI). Schools that are able to complete all the prerequisites before the course can offer Calculus BC. Calculus BC is a full-year course in the calculus of functions of a single variable. It includes all topics covered in Calculus AB plus additional topics. The content of Calc BC is designed to qualify the student for placement and credit in a course that is one course beyond that granted for Calc AB .

Students in Calculus BC are given a separate exam from students in Calculus AB. The BC Exam asks questions that would require a deeper understanding of the material learned in the course and focuses on some target areas that the Calculus AB exam does not. Simply put, the Calc BC exam is more difficult than the Calculus AB exam. The structure of these exams, however is the same. Each exam is 3 hours and 15 minutes in length and is split up into two sections. Students have a certain time allotted to each section. The first section is a multiple choice and tests the students' knowledge in many different areas. The second section consists of free response questions. This part allows the student to show his or her ability to solve problems that require a more complex chain of reasoning.

For Chemistry AP there is only one course offered. This course is designed to be the equivalent of a general chemistry course usually taken during the first year of college. This allows the student to take classes freshman year of college that are normally taken in his or her second year. There are many differences between a college level general chemistry class and the usual first high school course in chemistry. Some of these are in the textbook, the topics covered, the emphasis put on chemical calculations and formulations of principles, the kind of lab work done, the number of topics covered, and the time spent on the course by the student. There are some prerequisites for the enrollment in an AP Chemistry course. A student must have successfully completed the first course in high school chemistry and also completed a second year algebra course.

The structure of the AP Chemistry exam is similar to that of the AP Math exam. The AP Chemistry exam is 3 hours long and is also split up into two sections. Like math, the students are allowed a certain amount of time to complete each section. The first section contains many multiple choice questions. The second section is a free response section with qualitative problems, essays, and some chemical reaction writing.

The College Board offers two courses in AP Computer Science (AP CS). The first course, Computer Science A emphasizes program methodology. This course is equivalent to a first semester college course in CS. The second course is Computer Science AB and includes all topics covered in Computer Science A plus a more formal and in depth study of algorithms, data structures, and abstractions. Unlike AP Math and AP Chemistry, there are no prerequisites required for AP CS. This course serves as the students' first course in computer science.

Like the AP Math exam, there are two AP CS exams offered, one for each course. The CS AB exam covers more material and requires a better knowledge if the topics covered in the
course than does the CS A exam. Both AP CS exams are similar to the AP Math and AP Chemistry exams in that they both last 3 hours and are split up into two sections. Students must finish each section in a given amount of time. The first section consists of multiple choice questions. The second section contains free response questions that test a student's reasoning abilities. Both exams offered in AP CS test a student's ability to design, write, analyze, and document programs and subprograms.

For all subjects, each AP exam is scored out of 5,5 being the highest and 1 being the lowest. Many colleges recognize the AP exams, and reward students who do well on the exams. Generally students who receive a 3 or better are rewarded. Rewards often include college credit or grants of some kind. The credit may mean that the student will have placed out of a particular class, and will start ahead of the rest of the incoming students. For example an AP math student coming into WPI will start in Calc III and earn credit for Calc I and II. So a student who receives a 3 or better and starts in an advanced course will end up saving time and possibly money by not taking Calc I or II.

## Grade Point Average

In setting up an assessment scheme, one problem we came across was the fact that WPI has no official grade point average. The normal way to compute a GPA would to calculate it out of 4 , with an " $A$ " being 4 , " $B$ " being 3 , " $C$ " being 2 and a failing grade being 0 . The fact that WPI has no failing grades, forced us to come up with our own ways of computing the GPA. WPI gives a grade of "NR" if a student does not complete a course or fails to receive enough points to earn a "C". But this mark the student receives does not necessarily mean the student failed the class. This is why we aren't figuring an "NR" (no record) as a 0 , in the calculation.

Students receive "NR" in a class for many different reasons. Some students receive "NR" in a course because they can't handle the material in the class and end up with a failing grade, this would be a reason to count the "NR" as a 0 . But others "NR" classes because they can't find time to fit the class into their schedule and they end up dropping the class half way through the term. A student may do this because he or she knows the NR doesn't count against them, as a failed class. This would be a case where the student didn't deserve a 0 in the calculation of the grade average. No matter what the situation the student may take the class over again at a later date. This is why we had to account for both cases in our GPA calculations.

So we calculated two separate grade point averages, GPA and GPA II. GPA is the first way we calculated the average. This is the calculation that doesn't count an "NR" against a student. GPA, counts an "A" as 4 points, " B " as 3 , and " C " as 2 . GPA II, the second computation of the average, does count the "NR" against a student. GPA II, completes the average exactly the same as above except an "NR" is equal to 0 points. GPA II adds to the denominator of the average while adding nothing to the numerator, for this reason it will be smaller than the first GPA in all cases involving NR's.

## Microsoft Excel

All of our data and calculations are in an excel spreadsheet so we decided it would be best to briefly describe the program. This will help the reader to better understand some of the things we explain in the Method. Microsoft Excel is a spreadsheet program that is used to manipulate and sort and numerically analyze data. Using Excel allows you to easily analyze data and charts from various aspects of your spreadsheet. Basically Excel provides an easy way to manipulate a robust amount of data into something readable and presentable.

The Excel spreadsheet is broken down into rows (across) and columns (down). The spot at which each row and column intersect is called a cell. Each cell can hold one piece of your data. The data could be read in columns, like ours was, or in rows. The spreadsheet is like a table with the title of the columns up top, and maybe a counter down the side, and the cells in the middle our filled with data. In our spreadsheet the data was read down, starting at the top with the titles of each individual column and then read down from there.

In addition to holding just data, the cells can hold formulas. The formulas can do many things, and the result of the formula will be output into the particular cell the formula belongs to. There are many formulas in Microsoft Excel, the ones we used most include: sum, if, countif, and average. To use a formula you enter it into a cell beginning with the " $=$ " sign, an example is =sum(b2:e2). "Sum" is a formula that simply sums up the number in the range you input, and outputs the sum in the cell. If we take the example above, Excel is going to sum up the data in cells b2 to e2, and the answer will be output in the cell in which you input the formula. "If" is a logical formula that returns one value if the logic test is true and a second if the test is false. An example is $=\mathrm{if}(\mathrm{a} 2=\mathrm{b} 2,1,0)$, which would return a 1 if the value in cell " a 2 " was the same as in " b 2 " and would return a 0 if not. "Average" returns the average of the range of cells you input. And "countif" counts the amount of an argument in a range of cells, $=\operatorname{countif}(a 2: d 2,4)$ would
count the number of 4's in the range. Formulas make it very easy to manipulate and analyze data, especially when you have such a robust amount of data.

Microsoft Excel also has a programming aspect to it. The programming language you use is Visual Basic. Visual Basic is run along side Excel and can help you run huge projects in a short amount of time. The steps we used most in Visual Basic was running "for" loops to help us do tasks quickly. Without the loops we might have to copy and paste numerous times, which is not only tedious but very time consuming. A "for" loop is a loop that does something until a counter has run out. So we knew we needed copying and pasting done about 29,438 times, we had the loop run from $t=0$ to 29,438 , and every time the loop went through it copied another group data to where we wanted it. Basically Visual Basic is a programming language that can help you manipulate data further and much quicker.

## Main Hypothesis

In learning about the AP Program, we feel that AP students are, in fact, more able than non-AP students to perform serious college level work. Before we get into any analysis comparing AP students with non-AP students, we should state what we expect to find from our analysis. When we are finished with this project, we think our findings will show that the AP students perform better at the college level. Our data and analysis will either support or disprove this hypothesis.

There are many reasons as to why we think the analysis and comparison portion of our project will back up our hypothesis. The main one being AP courses offered in High School expose the students to college level work earlier than students who do not take AP level classes. Also through research of the AP program we have seen studies done comparing the performance of AP students vs. non-AP students in second level classes. The AP students having not taken the first level class at the college level, but rather in High School. While, the non-AP students have taken the first level course in college.

Some of our results, however, may not be as accurate as we would like. The reason for this could be that for some of the courses there were not enough students to get an accurate picture for comparison of AP students with non-AP students. For example, when we totaled up the number of AP Computer Science students from one complete class of students at WPI, there were only 15 . Compare this to the total number of AP Math students, which we found to be 132 . This is a huge difference when comparing these numbers to the whole population of students which is about 588. The population of AP Math students should not propose a problem because we are getting a good cross section of the population, but the group of 15 CS students is
something we should watch for and mention in our analysis. To get an accurate sample, most statistical tests like to see a sample size of 30 or more.

## METHOD:

The first thing we need to do is figure out who, at WPI, has taken Advanced Placement courses and has passed any AP exams. In order to do this we have to find out what courses taken by freshmen constitute a student who was involved in the Advanced Placement program, and identify or "tag" these students. This "tagging" of students will be different for each type of AP course taken by the students. For math, an "AP student" will be recognized as taking MA 1023(Calculus 3) or higher as his or her first math course at WPI. Also a student taking CH 1020, CH102X or higher as a first chemistry class is considered an AP student in that subject. A student taking CS 2005 as a first Computer Science class is considered AP in computer science.

## Identifiers

Identifying the data in the Excel file was one of the more difficult tasks we encountered thus far. Identifying the math AP students was what we did first. To do this we created an extra column in the spreadsheet labeled AP Math. This column consisted of a nested if then statement. The formula checked if it was the first semester for the student, if it was A term, and if the class was MA1003 or MA1023. If all this criteria was met the student received an "apm" in this column. The same was done for chemistry and computer science. The chemistry AP students were distinguished as students who took CH 1020 or CH 102 X as their first chemistry class. The problem with this was that we couldn't simply check if the student took CH1020 in A term of their freshman year, because some students held off and didn't take Chem II until B or C term. So what we did was check if the students had taken CH 1010 , if they had, they were taken out of contention. Next we checked the remainder of the students, the ones who hadn't taken Chem I, and checked if they had taken CH 1020 at any point their freshman year. If they had, they were identified with an "apch". Computer Science was done similarly to Chemistry. We had to first check if the student had taken CS1005, if they had, they were not AP Computer Science
students. The rest of the students were then tagged with an "apcs" if they had taken a 2000 level or higher CS class.

The identifier: "apch", "apm", "apcs" however was only placed in one row, the same row as the identifying class was. This proposed a problem to us, because what we eventually wanted to do was copy and paste the student's whole file to a separate page. To do this we needed an indicator that would tell us that all the rows in this student's file belonged on the AP page. To identify all the rows in the students file, we ran a macro. The macro consisted of simple "if, then" statements and a few "for loops". What the macro did was look for an "apm", when it found it, the corresponding student ID number was copied and pasted into another column. After this was done, another if then statement and for loop was run, this loop searched for any student ID that matched any of the ID's that had already been copied into the new column. If the loop found any that did match it placed an "apm" in a column next to the student ID. What was good about this was that it placed the "apm" next to any ID number that matched. So now we had every file of an AP student tagged from top to bottom. A sample of the code we used for the AP chemistry students is below. The apostrophes denote comments that explain each line of code:

$$
\begin{aligned}
& \mathrm{n}=0 \quad \text { 'sets } \mathrm{n} \text { (a counter) to zero } \\
& \text { For } \mathrm{i}=0 \text { To } 29440 \text { '" beginning of the first for loop } \\
& \text { If Cells }(3+\mathrm{i}, 28)=\text { "apcs" } \\
& \text { Then } \mathrm{n}=\mathrm{n}+1 \quad \text { ''if then loop that counts the amount of apcs students } \\
& \text { Next }
\end{aligned}
$$

For $k=0$ To $29440 \quad$ ' 2 nd for loop If Cells $(3+k, 28)=$ "apcs" Then ''if then statement

```
Cells(3+k, 13).Copy ''if the if is true, copies the id number
Cells(3 + j, 29).Select
ActiveSheet.Paste
Application.CutCopyMode = False
j=j + 1
End If
```

Next

```
For m = 0 To 29437 ''3 'rd for loop
    Forl=0 To n ''nested for loop
        If Cells(3 + m, 13) = Cells( 3 + 1, 29) Then ''checks if any of the id
                                    numbers match the new
                                    column of them
            Cells(3+m,30) = "apcs" ''places an "apcs" next to all the matching id
            End If
    Next
    Next
```

What we did next was run a macro that first searched for the "apm" in the new column, and if it found one, the macro would copy the needed information to another page. This information we copied to the new page was the year, term, class, grade, and student ID. This new sheet now consisted of only AP Math students. Similar sheets have been created for AP computer science and AP chemistry students. This now allowed us to not have to deal with the spreadsheet as a whole. We had each separate category of AP students in their own spreadsheet, and we could manipulate this data as much as we wanted without dealing with a whole population.

There were 132 AP math students, 31 AP chemistry students, and 15 AP computer science. This gives us a robust population of AP students to look at, out of the 588 total students. The only thing we need to look out for is if the differences in our GPA's in our analysis will be significant. There shouldn't be a problem with the math students, because the data contains a
good amount of students, but with the chemistry and computer science students there might be a problem. The problem could arise because 33 or 15 students is not a good proportion of the whole class. It could just so happen that the 15 students we took as AP computer science students, are the best or worst students in this particular class. This is just something that has to be watched for when we are doing our analysis.

## AP MATH

We began to manipulate the math data even further. The next step was to look at the AP math freshmen class only, so we ran a macro, similar to the above, that searched for first year students and placed their records in a separate table. The same was done for math classes only. We did this so we can analyze how the AP students did during their freshman year, how they did in math classes throughout the four years, and how they did in math classes during the freshman year only. We also separated AP Math students grades in MA2051 (Differential Equations) from all non-AP students. We did this because MA2051 relies heavily on the prior knowledge of Calculus I and II. The AP students did not take these courses at WPI and all the non-AP students did. This should be a good indicator in how well the AP courses in high school prepared these students for college level courses. All these figures that we come up with can be easily compared with the figures we receive from the population as a whole, and the non-AP students.

Now that we have the AP Math students arranged the way we want them to be:
I. AP Math students all years all classes.
II. AP Math students all years just math classes.
III. AP Math students freshman year all classes.
IV. AP Math students freshman year just math classes.
V. AP Math students MA2051 grades.
VI. Non-AP Math students MA2051 grades.

We can begin to gather data and compare all of these categories to the same categories of non AP students. The first data obtained was the number of A's, B's, C's, and NR's each student population had earned. To do this we had to use a counter command in the excel spreadsheet. The command $=$ COUNTIF(C4:C600,"A") (where C4:C600 is the range scanned, and the command will count the cell if there is an "A" in it) was used to calculate the number of A's obtained by the student population. The same command was also used to obtain the number of B's, C's, and NR's. From this we can also calculate the percentage of A's, B's, C's, and NR's obtained by each student population. To do this all we did was add up all the grades which gave us the total number of classes taken and put the number of A's received over this to find the percentage of A's earned. We also did this with the B's, C's, and NR's. This gives us some valuable data in which we can compare each group's performance at WPI.

## AP CHEMISTRY

A similar method was used in gathering data for the AP Chemistry students. Macros were run to organize our data the way we wanted to analyze them. The categories we ended up with to analyze were:
I. AP Chemistry students all years, all classes.
II. AP Chemistry students all years, just chemistry classes.
III. AP Chemistry students freshman year, all classes.
IV. AP Chemistry students freshman year, chemistry classes.

These were each organized into the spreadsheet and will be compared to the same categories of non-AP Chemistry students respectively. We then could calculate the number of A's, B's, C's, and NR's obtained by the students by using the same counter command we used earlier (=COUNTIF(C4:C600,"A") (where C4:C600 is the range scanned, and the command will count
the cell if there is an " A " in it)). We then calculated the percents of each grade by dividing the number of the grade being counted divided by the total number of grades. This data will help us in the comparison of AP Chemistry students from non-AP Chemistry students.

## AP COMPUTER SCIENCE

The data from AP Computer Science students were also gathered in a similar way. We used macros to organize the spreadsheet into the categories we wanted to analyze. After running a few programs we came out with the categories:
I. AP Computer Science students all years, all classes.
II. AP Computer Science students all years, Computer Science courses.
III. AP Computer Science students freshman year, all classes.
IV. AP Computer Science students freshman year, Computer Science classes.

These will be the categories analyzed with the same categories for non-AP Computer Science students respectively. Once we have all the different categories separated we calculated the total number of each grade. This was done by using the same counter command used above. Then the percentages were calculated by dividing the counted number of the grade percentage you are looking for divided by the total number of grades. We can now make comparisons of the performance of AP Computer Science students versus non-AP Computer Science students.

## GPA I and GPA II

Another important piece of data that could help us analyze the AP students performance versus non-AP students is the students GPA. This was easily done on the excel spreadsheet in a few steps. First we made a new column in the spreadsheet called GPA. We then used some "if, then" statements to turn the grades earned by the students into numbers that we could calculate. By using the equation $=\operatorname{IF}\left(\mathrm{F} 4=" \mathrm{~A} ", 4, \mathrm{IF}\left(\mathrm{F} 4=" \mathrm{~B} ", 3, \mathrm{IF}\left(\mathrm{F} 4={ }^{\prime} \mathrm{C} ", 2, \mathrm{FALSE}\right)\right)\right.$ ), this will designate all the A's received by the students as the number 4, B's are 3, C's are 2, and NR's are "False"
which means they are not used in calculating the GPA and does not count against the student. The second method of calculating a GPA for the students we designated as GPA II. This calculates the GPA in a slightly different way. First we created another column in the spreadsheet and called it GPA II. We then used another "if, then" statement to turn the actual grades into numbers that we could work with. By using the equation $=\mathrm{IF}\left(\mathrm{F} 4=\right.$ " $\mathrm{A} ", 4, \mathrm{IF}\left(\mathrm{F} 4=" \mathrm{~B} ", 3, \mathrm{IF}\left(\mathrm{F} 4=\mathrm{C} \mathrm{C}^{\prime}, 2, \mathrm{IF}(\mathrm{F} 4=\right.\right.$ "NR", $\left.\left.0, \mathrm{FALSE})\right)\right)$ ), every A recorded will be a 4 , B will be a 3 , C will be a 2 , and NR will count as a zero. In this case if the students receives a NR in a class it will count against them.

## Hypothesis Testing

A hypothesis test is a tool used in statistics to verify if results found in your data is statistically significant or just chance. A hypothesis test is a type of statistical inference. Statistical inference is the use of a subset of a population, sample, to draw conclusions on the whole population from which the sample is taken. Statistical inference is a method that relies on the distribution of the entire data set, and places emphasis on the normal distribution.

Hypothesis testing has 5 different components: scientific hypothesis, statistical model, statistical hypothesis, test statistic, and p-value. The scientific hypothesis is the hypothesis that you are trying to test through the experiment. The statistical model is the model that describes the data set. Statistical hypothesis are the two hypothesis (null hypothesis and alternative hypothesis) you are trying to prove or disprove, you must choose between them. One of the statistical hypothesis, the alternative one, will be the scientific hypothesis stated in terms of the statistical model. The test statistic is used to decide between the null hypothesis $\left(\mathrm{H}_{0}\right)$ and the alt0ernative hypothesis $\left(\mathrm{H}_{\mathrm{a}}\right)$. The last piece of a hypothesis test is the p -value, this value quantifies the relevance of your decision between $\mathrm{H}_{0}$ and $\mathrm{H}_{a}$.

The specific type of hypothesis test that we ran is called a one-sample hypothesis test for a population proportion. In this type of testing the data set is Y , which counts the number of successes in a certain number of binomial trials ${ }^{1}$. The number of trials, $n$, is known in the population proportion hypothesis test. What you are actually testing will be a hypothesis about p , which is the probability of a success. The sample size, n , will be large, greater than 30 , and the probability, p, will be reasonable, between .25 and .75 . With both these being true the

[^0]sample can be approximated, using the central limit theorem, by the normal distribution ${ }^{1}$. Being able to approximate the data with the normal distribution is very advantageous. The reason for this being, the normal distribution can be standardized and values can be received from tables, which can not be done so easily with the binomial distribution. Standardizing a normal random variable is an action that creates a new distribution with the same shape, but with mean at 0 and standard deviation of 1 . The test statistic in this case will be a standardized normal statistic $\mathrm{Z} . \mathrm{Z}$ will be equal to Y , our observed number of successes, minus the mean of the distribution, all divided by the standard deviation.

Statistical significance results when the amount of evidence in the data, against the null hypothesis and for the alternative hypothesis, is sufficient enough to reject the null hypothesis. However this can not tell us whether or not any difference we have found is of practical importance, like in our case. Statistical significance would mean there actually is a difference between the two results.

[^1]
## Conclusions Drawn from Analysis

Before we get into the conclusions that we drew from the results we should explain a few things. First, the category "other", corresponds to classes that students do not have to take and do get credit for them. The school does not give grades for this, so it was impossible for us to put a label on it either. Also when looking at the tables, the n value is simply the number of students who are in each category.

When examining the strength of AP Exams, the methodology used to compare AP students with non_AP, compares the performance in a second college course between those who have taken the first course in the college and the performance of AP students who were allowed to skip the first course. In almost all courses examined AP students had higher grades than those achieved by students without AP experience. This is what we essentially did with MA 1024 and MA 2501, which you will see later on in the report. While this result is important for colleges in that it helps them decide where AP course credit should be awarded, it falls short for other purposes. It is not as helpful as it could be to the individual college student, because one cannot know how well the AP students would have done had they taken the preliminary course instead of skipping straight to the second course. This is because AP students tend, on average, to display higher academic performance than do non-AP students. It also is not helpful to high school students who are trying to decide on the prospective value of taking an AP course.

From our analysis, we have come up with a number of different conclusions. In each separate subject we looked at every possible comparison that would be relevant to our study. We compared areas, for example, such as freshman year AP math students and freshman year nonAP math students. We also compared areas such as AP math students for four years with nonAP math students for four years as well as many other areas, including math classes only, math
classes during freshman year, etc. Comparisons similar to these were done for the two remaining subjects as well.

Our conclusions were drawn from the results we collected in the areas of GPA's. Both methods, GPA I and GPA II, were analyzed and compared. We also looked at the grade distribution in each of these categories. All these comparisons and analysis were done in order to accurately come up with a conclusion about whether AP students perform better than non-AP students.

When comparing these different subject areas we had to keep in mind that some areas do not have as many students in the AP program as the others. Due to this fact some of our analysis looks to be not as accurate as we would have liked. For example, there were only 15 AP computer science students as opposed to 132 AP math students. Both of these numbers are out of the 588 total students. So in this case, a more accurate analysis and conclusion would be drawn in the AP math.

## MATH

When analyzing and comparing the different categories in math, we found interesting results. There were many instances were there was a substantial difference in performances of AP students with non-AP students. Likewise, we found other times when there were close to no differences in their performances. In light of this we will go through and explain these differences as being relevant to our study. When reading these results, keep in mind that the number of AP Math students was 132 as compared to 456 non-AP Math students.

Table 1: AP Only vs. All Students

| All Students; All classes ( $n=588$ ) |  |  | AP Math; All classes$(n=132)$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | first year | four years |  | first year | Four years |
| \%A | 23.65\% | 32.12\% | \%A | 30.38\% | 37.27\% |
| \%B | 28.92\% | 29.89\% | \%B | 38.61\% | 29.29\% |
| \%C | 22.10\% | 19.35\% | \%C | 13.29\% | 14.73\% |
| \%NR | 16.70\% | 13.83\% | \%NR | 8.23\% | 11.33\% |
| \%other | 8.62\% | 4.81\% | \%other | 9.49\% | 7.38\% |
| GPA | 3.020715 | 3.156981 | GPA | 3.207692 | 3.27717 |
| GPA II | 2.468597 | 2.698451 | GPA II | 2.916084 | 2.87644 |

We did the hypothesis test in our project to verify if the percentage of NR's in the AP math population, .1133 , was significantly different then the percentage of NR's in the whole population, .1383 . There is obviously a difference in the 2 percentages but we ran the test to see if it was by random chance or if the numbers were actually different.

In our case $Y$, the number of successes in the entire population, was the number of NR's out of the $n$, which is 132 . So $p_{o}$ is equal to the probability of getting an NR which is .1383 . So $Y$ follows a binomial distribution with parameters $\mathrm{n}=132$ and $\mathrm{p}_{0}=.1383$. We then approximate Y using the normal distribution, we can do this with such a low $\mathrm{p}, .1383$, because our n is so large. And we are looking for the probability that Y is greater than or equal to 14.95. The reason we pick 14.95 is because it is the number of NR's in the sample population (.1133*132). Which is equal to one minus the probability that Y is less than 14.95. The reason we would rather look for the probability that Y is less than some number, is because finding probabilities that Y is less than some number can be looked up from a table. The next step is to standardize the random variable, we do this because the probabilities are easily found in the table. So we subtract 18.26 , which is the number of NR's in the entire population, from 14.95 and then divide by the standard deviation, which is the square root of $\mathrm{n}^{*} \mathrm{p}_{0} *\left(1-\mathrm{p}_{\mathrm{o}}\right)$. The solution of this comes out to be -.834549 . And this new standardized random variable is Z . We are now looking for one minus the
probability that Z is less than .834549 . We can obtain this value directly from the table. What we come up with is a p-value of $.2083{ }^{1}$.

The following results are from students at WPI and are taken from all their classes freshman year, not just math. There is a rather substantial difference in the GPA I between AP Math and Non-AP Math. (recall the GPA I does not count NR's against the students). So for AP Math the GPA I is 3.21 and for non-AP Math it is 2.97 . For GPA II, when you count the NR's as 0 , AP Math students had a 2.92 GPA II average and for non-AP math it was 2.39 . So in both cases, the AP Math students had better averages than non-AP Math.

Table 2: AP vs. Non-AP Only

| AP Math; All classes ( $\mathrm{n}=132$ ) |  |  | All Students less AP Math: All classes ( $\mathrm{n}=456$ ) |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | first year | Four years |  | first year | four years |
| \%A | 30.38\% | 37.27\% | \%A | 22.37\% | 30.44\% |
| \%B | 38.61\% | 29.29\% | \%B | 29.07\% | 30.10\% |
| \%C | 13.29\% | 14.73\% | \%C | 24.29\% | 20.81\% |
| \%NR | 8.23\% | 11.33\% | \%NR | 18.36\% | 14.63\% |
| \%other | 9.49\% | 7.38\% | \%other | 5.91\% | 4.03\% |
| GPA | 3.207692 | 3.27717 | GPA | 2.974661 | 3.118371 |
| GPA II | 2.916084 | 2.87644 | GPA II | 2.394308 | 2.643038 |

In Table 2 a hypothesis test was done to see if the difference between the GPA II averages was significant. The GPA II averages we are looking at are for freshman year, all classes. The pvalue was .0113 . This gives us strong evidence against $\mathrm{H}_{0}$, which was the two averages were the same, and strong evidence for the alternative hypothesis; the AP math population has a statistically significant higher GPA II than the non-AP group.

When comparing the non-AP students with all students, there was a very small difference, if any, in the percentages of A's, B's, C's and NR's received by non-AP students vs. all students.

[^2]Table 3: All Results - Math Classes Only

All Students less AP Math ( $\mathrm{n}=456$ )

|  | first year | Four <br> years |
| :--- | ---: | ---: |
| \%A | $14.50 \%$ | $13.31 \%$ |
| \%B | $28.39 \%$ | $27.53 \%$ |
| \%C | $31.92 \%$ | $33.00 \%$ |
| \%NR | $17.59 \%$ | $20.13 \%$ |
| \%other | $7.60 \%$ | $6.04 \%$ |
| GPA | 2.76705 | 2.7333 |
| GPA II | 2.24031 | 2.14766 |

AP math ( $\mathrm{n}=132$ )

|  | First | four |
| :--- | ---: | ---: |
|  | year | years |
| \%A | $20.17 \%$ | $21.93 \%$ |
| \%B | $25.46 \%$ | $27.72 \%$ |
| \%C | $14.31 \%$ | $16.45 \%$ |
| \%NR | $7.01 \%$ | $9.27 \%$ |
| \%other | $33.05 \%$ | $24.63 \%$ |
| GPA | 3.1 | 3.08 |
| GPA II | 2.7735 | 2.703 |

All Students ( $\mathrm{n}=588$ )

|  | first year | four |
| :--- | ---: | ---: |
|  |  | years |
| \%A | $16.26 \%$ | $15.55 \%$ |
| \%B | $27.47 \%$ | $27.51 \%$ |
| \%C | $26.96 \%$ | $28.92 \%$ |
| \%NR | $14.62 \%$ | $17.46 \%$ |
| \%other | $14.69 \%$ | $10.57 \%$ |
| GPA | 2.84867 | 2.81422 |
| GPA II | 2.36059 | 2.26478 |

In the next part we made the same comparisons but were concerned only with math classes that students took (Table 3). The results were similar to those we saw above when we used all classes. The AP Math students had a much better percentages of A's, B's, C's, NR's than did non-AP Math. Likewise, the GPA I and GPA II were also better for the AP Math students. And when we compare the non-AP students with all students we found that there was not much of a difference in percentages and GPA's, as was the case above with all classes.

## Differential Equations

Now the next thing we looked at was look at how both groups performed in differential equations, MA 2051 (Table 4).

Table 4: MA 2051

All Math AP; 2051 ( $\mathrm{n}=132$ ) All Students; MA
2051 ( $n=588$ )

|  | MA |  | MA 2051 | n=456) | MA 2051 |
| :--- | :--- | :--- | ---: | :--- | ---: |
|  | 2051 |  |  |  |  |
| \%A | $20.18 \%$ | \%A | $14.41 \%$ | \%A | $12.12 \%$ |
| \%B | $35.09 \%$ | \%B | $28.60 \%$ | \%B | $26.72 \%$ |
| \%C | $34.21 \%$ | \%C | $35.07 \%$ | \%C | $35.54 \%$ |
| \%NR | $10.53 \%$ | \%NR | $20.46 \%$ | \%NR | $23.69 \%$ |
| \%other | $0.00 \%$ | \%other | $1.46 \%$ | \%other | $1.93 \%$ |
| GPA | 2.5625 | GPA | 2.735294 | GPA | 2.685185 |
| GPA II | 2.54386 | GPA II | 2.167373 | GPA II | 2.036517 |

[^3]The GPA here was computed from all the grades in just MA 2051, differential equations.
For MA 2051 we found that AP Math students did better overall than non-AP. When comparing the GPA's we found interesting results. The GPA I for non-AP was found to be higher than AP. Non-AP GPA I was 2.69 and for AP it was 2.56. However the GPA II yielded different results. The AP students had a 2.54 and non-AP had a 2.04 GPA II average. This is a big difference and is discussed in the summary. And as was done before, we compared the nonAP with all students. In MA 2051 we found the same type of thing as we found when comparing the other sections. There was almost no difference in any of the grade percentages as well as the GPA's.

A hypothesis test was run to check for significant differences in the percentage of NR's in the non-AP and AP groups, for all grades in the MA 2051 class. This time around we must be careful because these are two separate types of populations. In other words, one group is not a sample or subset of the other. The AP math students had a proportion of NR's equal to .1053 and the non-AP students' proportion was .2369 . Looking at these two numbers you would think there would be a significant difference. And that is exactly what the hypothesis test told us. The p-value we obtained was extremely small, which means there is very strong evidence against the two values being the same and for the two values being different. In conclusion, the AP math percentage of NR's in MA 2051 is less than the percentage of NR's for non-AP students in the same class. And this fact is backed up by a very low p-value.

## Calculus IV

In MA 1024 we found major differences which could mean a lot in the final scheme (Table 5). In reference to the GPA I and II, the AP students received a much higher average. The GPA II average for AP was 3.08 and for non-AP it was 2.18. As this comparison shows there is a huge difference in the performances of AP and non-AP Math students. However we found that all
students combined had a slightly higher percentage of A's, B's and GPA II's. All other categories, C's and NR's and GPA I, were the same. This result is due to the fact that AP students did so much better than non-AP.

Table 5: MA 1024

| All Students; MA |  |
| :--- | ---: |
| 1024 $(\mathrm{n}=588)$ |  |
| MA 1024 |  |
| \%A | $20.00 \%$ |
| \%B | $31.68 \%$ |
| \%C | $27.43 \%$ |
| \%NR | $19.47 \%$ |
| \%other | $1.42 \%$ |
| GPA | 2.90604 |
| GPA II | 2.332136 |


| All Students less AP math; |  | All Math AP; MA 1024 ( $n=132$ ) |  |
| :---: | :---: | :---: | :---: |
|  | MA 1024 |  | MA |
|  |  |  | 1024 |
| \%A | 19.92\% | \%A | 35.25\% |
| \%B | 25.87\% | \%B | 47.54\% |
| \%C | 29.36\% | \%C | 9.84\% |
| \%NR | 23.82\% | \%NR | 5.74\% |
| \%other | 1.03\% | \%other | 1.64\% |
| GPA | 2.874317 | GPA | 3.27434 |
| GPA II | 2.182573 | GPA II | 3.08333 |

Once again we ran a hypothesis test on the two means to verify if the difference is significant or not. We obtained a test statistic, Z of 2.133 , which yielded a p -value of .0166 . If a p-value is less than .01 , the evidence against the two values being the same is very strong. Our p-value in this case is not too much greater than that value, so our evidence for the two means being different is strong. This means there is a statistically significant difference in the percentage of NR's in MA 1024 between the AP math students and the non-AP math students.

It is important to see how each group did in both these classses, MA1024 and MA 2051. This is because both non-AP and AP Math students coming into WPI must take both of these classes. However AP students, because they started in Calc III, did not take Calc I or II at WPI, whereas non-AP Math students did take Calc I and II. So it will be important to see how these students compare in each class, MA 2051 and MA 1024, with one group having taken Calc I and II at WPI and the other taking the equivalent of Calc I and II at their high schools. Another thing to keep in mind when looking at the results is the fact that in MA 1024 AP and non-AP Math students were not in the same class. This is because AP students would take MA 1024 in B-term
of freshman year, whereas non-AP students took the same class in D-term of freshman year. So this would mean that a professor may teach a class of only AP Math students in B-term a slight bit different than he or she would had the class been a mixture of AP and non-AP. This is why it is also important to include MA 2051 because this class is a mixture of AP and non-AP that are from different backgrounds.

## CHEMISTRY

For chemistry we made comparisons that were similar to those done in math. Some of the results for chemistry, however, were a lot different than what we got in math. We found more cases in chemistry where the non-AP students did better than the AP students. Bear in mind that the number of chemistry students is only 31 .

## Table 6: All Classes

| All students ( $\mathrm{n}=588$ ) |  |  | AP Chem ( $\mathrm{n}=31$ ) |  |  | All Students less AP chem ( $\mathrm{n}=557$ ) |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | First year | four years |  | first year | four years |  | first year | four years |
| \%A | 23.65\% | 32.12\% | \%A | 22.68\% | 31.05\% | \%A | 23.71\% | 32.18\% |
| \%B | 28.92\% | 29.89\% | \%B | 25.27\% | 29.82\% | \%B | 29.12\% | 29.90\% |
| \%C | 22.10\% | 19.35\% | \%C | 24.62\% | 20.25\% | \%C | 21.96\% | 19.30\% |
| \%NR | 16.70\% | 13.83\% | \%NR | 16.20\% | 13.20\% | \%NR | 16.73\% | 13.86\% |
| \%other | 8.62\% | 4.81\% | \%other | 11.23\% | 5.69\% | \%other | 8.48\% | 4.76\% |
| GPA | 3.02071 | 3.15698 | GPA | 2.97321 | 3.13317 | GPA | 3.02336 | 3.15831 |
| GPA II | 2.4686 | 2.69845 | GPA II | 2.43066 | 2.69479 | GPA II | 2.47071 | 2.69866 |

The first thing we did was compare freshman year only AP Chemistry students with freshman year only non-AP Chemistry students considering all classes taken. Unlike math, the non-AP Chem students did slightly better that AP Chem. There was less than a $3 \%$ difference in the percentages of A's, B's, C's, and NR's between non-AP and AP Chem students. In all cases the non-AP students did better than AP. Furthermore, the GPA I and II for non-AP students was .04 a better average than that of AP in both cases. So in this case, although only by a slight amount,
the non-AP Chemistry students performed better than AP Chemistry students. We then found the same results when comparing the two groups over all four years of college including all classes.

We then looked at freshman year only focusing on chem classes only and compared AP Chem and non-AP Chem. We found that non-AP students performed significantly better than non-AP. The percentage of A's and B's was $5 \%$ better for non-AP students than AP. The GPA I for nonAP students was 2.86 and the GPA I for AP students was 2.76 . Also the GPA II for non-AP was 2.53 and for AP was 2.49. Again non-AP students performed better. When we compared nonAP students with all students we found than they were almost the same. So with chemistry, we found that non-AP performed better than AP in most cases.

## COMPUTER SCIENCE

First, there are only 15 AP computer science students. We looked at Computer Science and found mixed results (Table 7). The percentage of A's was higher for AP and the percentage of B's was higher for non-AP. Also the GPA I and II were better for AP CS students.

Table 7: All Classes

| $\begin{aligned} & \text { All Stud } \\ & (\mathrm{n}=588) \end{aligned}$ |  |  |
| :---: | :---: | :---: |
|  | first year |  |
| \%A | 23.65\% | 32.12\% |
| \%B | 28.92\% | 29.89\% |
| \%C | 22.10\% | 19.35\% |
| \%NR | 16.70\% | 13.83\% |
| \%other | 8.62\% | 4.81\% |
| GPA | 3.02071 | 3.15698 |
| GPA II | 2.4686 | 2.69845 |


| AP Computer Science <br> ( $\mathbf{n}=15$ ( $)$ | first year four |  |
| :--- | ---: | ---: |
|  | years |  |
| \%A | $32.53 \%$ | $46.16 \%$ |
| \%B | $19.68 \%$ | $21.98 \%$ |
| \%C | $18.88 \%$ | $14.30 \%$ |
| \%NR | $15.66 \%$ | $10.92 \%$ |
| \%other | $13.25 \%$ | $6.63 \%$ |
| GPA | 3.19209 | 3.38644 |
| GPA II | 2.61574 | 2.99025 |


| All Students less AP CS <br> ( $\mathbf{n}=573$ ) | first year four |  |
| :--- | ---: | ---: |
|  | years |  |
| \%A | $23.40 \%$ | $31.75 \%$ |
| \%B | $29.19 \%$ | $30.11 \%$ |
| \%C | $22.20 \%$ | $19.49 \%$ |
| \%NR | $16.73 \%$ | $13.90 \%$ |
| \%other | $8.48 \%$ | $4.76 \%$ |
| GPA | 3.01604 | 3.15074 |
| GPA II | 2.46459 | 2.69078 |

We then compared the non-AP with AP looking at only computer science classes only (Table 8) for freshman year and found that AP had $42 \%$ A's whereas non-AP had $18 \%$. On the other hand the percentage of B's for non-AP was $32 \%$ and for AP it was $16 \%$. But we then found that, in both cases, the GPA I and II for AP CS was more than .05 points more than non-AP CS students. But when we compared non-AP students with all students we found that they were the almost the same, with the all students GPA I and II slightly better than non-AP.

Table 8: CS Classes Only

| All students <br> (n=588) | first year four |  |
| :--- | ---: | ---: |
|  | years |  |
|  | \%A | $19.51 \%$ |
| \% | $27.37 \%$ |  |
| \%B | $30.82 \%$ | $30.28 \%$ |
| \%C | $21.29 \%$ | $23.34 \%$ |
| \%NR | $21.95 \%$ | $17.18 \%$ |
| \%other | $6.43 \%$ | $1.84 \%$ |
| GPA | 2.97523 | 3.04973 |
| GPA II | 2.27725 | 2.51607 |


| All Students less AP CS <br> ( $\mathbf{n}=573$ ) |  |  |
| :--- | ---: | ---: |
|  | first year four |  |
|  | years |  |
| \%A | $17.86 \%$ | $25.71 \%$ |
| \%B | $31.90 \%$ | $30.87 \%$ |
| \%C | $22.14 \%$ | $24.18 \%$ |
| \%NR | $22.62 \%$ | $17.65 \%$ |
| \%other | $5.48 \%$ | $1.58 \%$ |
| GPA | 2.9404 | 3.01895 |
| GPA II | 2.23678 | 2.47745 |

AP CS ( $\mathrm{n}=15$ )

|  | first year four |  |
| :--- | ---: | ---: |
|  | years |  |
| \%A | $41.94 \%$ | $59.41 \%$ |
| \%B | $16.13 \%$ | $18.81 \%$ |
| \%C | $9.68 \%$ | $6.93 \%$ |
| \%NR | $12.90 \%$ | $7.92 \%$ |
| \%other | $19.35 \%$ | $6.93 \%$ |
| GPA | 3.47619 | 3.61628 |
| GPA II | 2.92 | 3.30851 |

## IQP AND MQP

In addition to checking the other categories we also checked the grade distributions and GPA's of the IQP's and MQP's. All students in our data set have completed both of the projects and we figured it would be interesting to analyze how AP students did on them versus non-AP.

Table 9: AP vs. Non-AP

## IQP and

MQP

|  | AP | Non AP |
| :--- | ---: | ---: |
| \%A | $71.27 \%$ | $64.52 \%$ |
| \%B | $18.29 \%$ | $25.09 \%$ |
| \%C | $6.78 \%$ | $6.77 \%$ |
| \%NR | $3.66 \%$ | $3.62 \%$ |
| GPA | 3.66948 | 3.59917 |
| GPA II | 3.53523 | 3.469 |

## SAT Data

Another statistic checked for were the SAT scores for the different populations (Table 10). The SAT's are another type of high school testing controlled by the College Board. So it would be interesting to see how the overall population did on their SAT's, we also checked how the AP math group did. We checked the overall SAT scores and how the group did on the math portion of the test. Also we figured out the standard deviation of each set of scores. The standard deviation of a set of data, is a measure of spread. The standard deviation is the mean distance of the data values from the mean of the data value.

The AP math mean SAT score was a bit higher than the all student average. The AP math average was 1216 for overall score, with the SAT math average was about 692 . Where the math average for all students was around 654. The standard deviation for the AP math is 61 which is lower than the all student which is 67 . Which makes total sense because the AP math is a smaller sample then the all student set. A lower standard deviation means that on average all the scores are closer to the mean, the data is more closely packed. In general this means that the highest all student SAT score should be higher than the highest AP math SAT score, and the lower should be lower.

Table 10: AP vs. All Students

## SAT Averages; All

 Students|  | All <br> students | AP Math |
| :--- | :--- | :--- |
| Total | 1172.238 | 1216.048 |
| Math | 654.6359 | 691.7742 |
| St. Dev. | 67.13242 | 61.12826 |

A hypothesis test was run to check for significant evidence that the AP math SAT average was higher than the all student average. Keep in mind that we only looked at the math section of the SAT, ignoring the verbal part. What we came up with was a p-value of .2676 ,
which means there is no evidence to say the mean SAT score for AP math is higher than 654 (the mean for all students' SAT score). Basically this means that the average SAT score for AP math is statistically the same as the average SAT score for the whole population.

The results we obtained from this hypothesis test on the SAT scores are very important to our project. The fact that the averages for AP math and the entire population are the same means that the AP math students did not actually perform better on the math portion of the SAT. And what this means is that the AP math students, in general, do not start of with a higher knowledge of mathematics when they take the AP courses. SAT's are usually taken towards the end of the junior year of high school while AP courses are taken through the senior year and tests are taken at the end of that year, May. Since the AP math students performed better in many categories than the non-AP and the whole population, one may conclude that the AP courses did actually provide these students with help.

## GENDER

One thing we did a quick check on was to see the distribution of gender between our populations. We wound up having a class of 588 that consisted of $78 \%$ male and $22 \%$ female. With the AP group, and AP chemistry group having a distribution very close to that, both about $77 \%$ to $23 \%$. The AP math was similarly close with $75 \%$ male and $25 \%$ female. But the distribution of gender for the AP CS turned out to be $93 \%$ male and $6 \%$ female. Considering there were only 15 AP CS students, and the fact that the majority of students in that major is male this number is not that alarming. If the numbers grew in that group the percentages would most likely converge to something closer to the distribution we had for the overall data set.

Table 11
Distribution of Gender

|  | Male | Female |
| :--- | ---: | ---: |
| Total | $78.44 \%$ | $21.56 \%$ |
| All AP | $77.44 \%$ | $22.56 \%$ |


| AP Math | $75.00 \%$ | $25.00 \%$ |
| :--- | ---: | ---: |
| AP Chem | $77.42 \%$ | $22.58 \%$ |
| AP CS | $93.33 \%$ | $6.67 \%$ |

## Summary:

The hypothesis we had going into this project was, AP students perform better in college than students who have not taken AP courses or exams. The robust amount of data we began with has been manipulated and rearranged down into something readable, and we have drawn many conclusions and results from this analysis. We will now summarize the results, and interpret them. This summary will interpret what all the tables and data has already told us in raw form. The summary will try and put words behind all the data that has been presented. Also we will state whether or not our hypothesis has been proven or disproven, or if their wasn't enough substantial evidence to decide either way.

We will interpret how each group of AP students, math, chemistry, computer science, did in each respective category. We compared the AP students in each subject with students who were not AP in that particular subject, and against the population as a whole (which included
them). We analyzed the performance of AP students and non-AP students in classes that corresponded to which subject the AP students were from (AP chemistry vs. non-AP chemistry in chemistry classes only). We investigated how these students did against each other just freshman year, and how they did in the particular subject just freshman year. We also went as far as to see how these students did in particular classes, MA 2051 and MA 1024 (differential equations and calculus IV).

One major problem we ran into in analyzing our data, was the number of AP students we had in each category. Overall we had 588 students to deal with, of which about 410 were nonAP and the remaining 178 were AP students. AP math was our most robust data set, consisting of 132 students, which is a large enough number to dampen any problems we should encounter that require n to be large. But our other two groups of AP students, CS and chemistry only accounted for 46 of the 178 AP students. Chemistry consisted of 31 students, which is relatively small compared to the size of the AP math group. This left only 15 AP computer science students. This proposes a problem to us because we are essentially comparing 573 non- AP students to only 15 . The reason this is a problem is because the 15 students could just happen to be the 15 top students in the class, or the 15 lowest students in the class, and there is not enough size in the group to bring up or down their averages. Whereas, in the AP math group we could have the 15 top students within the 132 , but with so many other students to flatten out the curve a little bit.

## Math

We focused a little more on the AP math population because it was so much larger than the other two groups of AP students. Within the AP math set of students, we looked at many different results, including: SAT math scores, performance in MA 1024 and 2051, performance
freshman year, and performance in math classes alone. Overall the AP math students performed better than the entire data set of students and better than the non-AP students.

Looking at the SAT scores of the AP math set vs. the SAT scores of the entire population allows us to obtain a picture of how these students did before they have taken any of the AP courses. The SAT's are another form of College Board tests, taken by high school students, before or in the beginning of their senior year. While the AP courses are taken during the senior year, and the AP exams are taken in May, after the senior year in high school. We found the AP math students to have an average of 692 on the SAT math, and the entire population had an average of 654. Although these numbers seem different, we found, through a hypothesis test, that the two numbers were not statistically significant at all. This means that the AP courses and the AP exams help these students learn greater amount of material, because the AP students perform better in most categories than the entire population.

Checking the performance in MA 2051 and MA 1024 was important because these are two math classes that AP math students and non-AP students would definitely take, either together or not, at WPI. In Calculus IV the two groups of students, AP and non-AP, may not necessarily be taking the class together. This is because the AP students are starting in Calc III and taking Calc IV during B-term most likely, while the non-AP students are not taking Calc IV until D-term of their freshman year. But the two groups of students could be taking MA 2051 during the same term, maybe A-term of the sophomore year for example. This is a factor in how the professors will teach the particular course, if the AP and non-AP are together the course will be designed differently than if the course consisted of only AP students. At any rate the AP students performed much better in MA 1024 than the other two groups, non-AP and the entire population. This is not as significant as the data in MA 2051, because it is unlikely that the non-

AP and AP students took MA 1024 together. In MA 2051, the AP students have a higher percentage of A's, B's and about the same percentage of C's as the other groups. As for NR percentage, the AP students had a much lower proportion of NR's than the other two. The hypothesis test we ran checking the NR percentage showed us that there is significant evidence proving that these two numbers are very different. This means MA 2051, a class that AP and non-AP students take together, will be passed by the AP students while the non-AP do not receive a passing grade.

We also looked at how these students did in math classes only. The AP students received a high percentage of "other" which just means the students got credit for the class, but did not have to complete the course (Calc I and II). The reason for the AP percentage of "other" is so high, is because these students don't take Calc I or Calc II, which automatically gives them an "other" in 2 courses. Overall the 132 AP math students performed much better in math classes only. In all the categories that we checked, we found the AP math students to perform better than the non-AP students.

## Chemistry

The AP chemistry data set only consisted of 31 students, which we were comparing to 557 non-AP students. We compared the two groups of students in categories similar to the ones we used in AP math, freshman year, chem classes only, and all classes. What we found was a little different from the AP math, the non-AP chemistry group did the same if not a little better than the AP chemistry.

When looking at chemistry classes only we found that all 3 groups, entire population, AP chemistry students, and non-AP chemistry students, did basically same. They have almost identical GPA's and grade distributions. Once again the AP population has a high percentage of "other" this is for the same reason that the AP math students had a high percentage of it. Overall
the it looks like the non-AP chemistry performed better than the AP chemistry, but the differences in the averages, and percentages are too small to be statistically significant. This means that the AP chem and the non-AP performed the same in our study.

## Computer Science

The number of students in the AP computer science set was only 15 . This could be because most students that come into WPI to pursue a major in Computer Science or something similar want a review of sorts, and start in a class that contains material they already know. Once again we looked at how these students did in computer science classes alone, freshman year, and in all classes.

The AP computer science students obtained an extremely high percentage of A's, which in turn yielded a much higher GPA than the other two groups. This is another case like the AP math that the AP students did much better than the rest of the group. The only thing with this group is that there is only 15 students to look at, this is against a set of 573 . This is a huge difference and could weigh some of the averages to one side or another, which could've happened in our study. Or not, the CS students could just perform that much better than the rest of the set of students.

## Conclusion

In conclusion, we were given a data set that contained a class of 588 students here at WPI. We analyzed and manipulated the data, and came up with results that would either prove or disprove our hypothesis about AP students doing better than non-AP students.

In math AP students performed consistently better than non-AP in most all categories. This includes MA 1024 as well as MA2051. We found that in Chemistry, AP and non-AP performed the same. For Computer Science, we found that the AP students did much better than non-AP. So overall we feel that AP students do perform better than non-AP students. This conclusion is consistent with our hypothesis given at the start of our project, which stated that AP students are more prepared for serious college-level work than non-AP students.

## References

Information on Advanced Placement Program. College Board; (www.collegeboard.org), 2000

Jennings, Roger (1997). Using Access 97 Second Edition. Indianapolis, IN: Que Corporation
Petruccelli, J., Nandram, B., Chen, M. (1999). Applied Statistics for Engineers and Scientists. Upper Saddle

River, NJ: Prentice Hall, Inc

Totals * Other are classes students get credit for, but haven't taken.
** (n) represents the number of students in the category

| All Students; All classes ( $\mathrm{n}=588$ ) |  |  |
| :---: | :---: | :---: |
|  | first year | four years |
| \%A | 23.65\% | 32.12\% |
| \%B | 28.92\% | 29.89\% |
| \%C | 22.10\% | 19.35\% |
| \%NR | 16.70\% | 13.83\% |
| \%other | 8.62\% | 4.81\% |
| GPA | 3.020715 | 3.156981 |
| GPA II | 2.468597 | 2.698451 |


| All students; less any AP $\quad(\mathrm{n}=410)$ |  |  |
| :--- | ---: | ---: |
|  | first year | four years |
| \%A | $22.35 \%$ | $30.42 \%$ |
| \%B | $29.53 \%$ | $30.27 \%$ |
| \%C | $24.22 \%$ | $20.76 \%$ |
| \%NR | $18.23 \%$ | $14.60 \%$ |
| \%other | $5.68 \%$ | $3.94 \%$ |
| GPA | 2.975455 | 3.118504 |
| GPA II | 2.400478 | 2.644484 |


| IQP and MQP |  |  |
| :--- | ---: | ---: |
| AP | Non AP |  |
| \%A | $71.27 \%$ | $64.52 \%$ |
| \%B | $18.29 \%$ | $25.09 \%$ |
| \%C | $6.78 \%$ | $6.77 \%$ |
| \%NR | $3.66 \%$ | $3.62 \%$ |
| GPA | 3.66948 | 3.59917 |
| GPA II | 3.53523 | 3.469 |


| Distribution of Gender |  |  |
| :--- | ---: | ---: |
|  |  |  |
| Male | Female |  |
| Total | $78.44 \%$ | $21.56 \%$ |
| All AP | $77.44 \%$ | $22.56 \%$ |
| AP Math | $75.00 \%$ | $25.00 \%$ |
| AP Chem | $77.42 \%$ | $22.58 \%$ |
| AP CS | $93.33 \%$ | $6.67 \%$ |

## Numbers of Each AP Subject

| Total | 588 |
| :--- | ---: |
| Total AP | 178 |
| Math | 132 |
| Chem | 31 |
| CS | 15 |

SAT Averages; All Students All students AP Math
Total $1172.238 \quad 1216.048$
Math $654.6359 \quad 691.7742$
St. Dev. 67.13242 61.12826

All Students；All classes first year


## All students；All classes four years



| 图A |
| :--- |
| 图B |
| $\square C$ |
| $\square N R$ |
| ■ other |

All students；less AP All Classes First year


图
图B
$\square C$ ［NR
图 other

All Students；Less AP All Classes Four years


| 图A |
| :---: |
| 园B |
| ［10 |
| TNR |
| 図 other |



IQP and MQP; Non AP


| $\square A$ |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |


| All Students; All classes ( $\mathrm{n}=588$ ) |  |  |
| :---: | :---: | :---: |
|  | first year | four years |
| \%A | 23.65\% | 32.12\% |
| \%B | 28.92\% | 29.89\% |
| \%C | 22.10\% | 19.35\% |
| \%NR | 16.70\% | 13.83\% |
| \%other | 8.62\% | 4.81\% |
| GPA | 3.020715 | 3.156981 |
| GPA II | 2.468597 | 2.698451 |

* Other are classes students get credit for, but haven't taken.
** ( $n$ ) represents the number of students in the category

| AP Math; All classes | $\begin{array}{l}\text { ( } \mathbf{n}=132 \text { ) }\end{array}$ |
| :--- | ---: | ---: |
|  | first year |
| four years |  |$]$| \%A | $30.38 \%$ | $37.27 \%$ |
| :--- | ---: | ---: |
| \%B | $38.61 \%$ | $29.29 \%$ |
| \%C | $13.29 \%$ | $14.73 \%$ |
| \%NR | $8.23 \%$ | $11.33 \%$ |
| \%other | $9.49 \%$ | $7.38 \%$ |
| GPA | 3.207692 | 3.27717 |
| GPA II | 2.916084 | 2.87644 |

All Students less AP Math ( $\mathrm{n}=456$ ) All classes

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $22.37 \%$ | $30.44 \%$ |
| \%B | $29.07 \%$ | $30.10 \%$ |
| \%C | $24.29 \%$ | $20.81 \%$ |
| \%NR | $18.36 \%$ | $14.63 \%$ |
| \%other | $5.91 \%$ | $4.03 \%$ |
| GPA | 2.974661 | 3.118371 |
| GPA II | 2.394308 | 2.643038 |

AP math; math classes only ( $\mathrm{n}=132$ )

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $20.17 \%$ | $21.93 \%$ |
| \%B | $25.46 \%$ | $27.72 \%$ |
| \%C | $14.31 \%$ | $16.45 \%$ |
| \%NR | $7.01 \%$ | $9.27 \%$ |
| \%other | $33.05 \%$ | $24.63 \%$ |
| GPA | 3.1 | 3.08 |
| GPA II | 2.773504 | 2.703 |

All Students less AP math;
MA 2051, 1024 ( $\mathrm{n}=456$ )
MA 2051
MA 1024

| \%A | $12.12 \%$ | $19.92 \%$ |
| :--- | ---: | ---: |
| \%B | $26.72 \%$ | $25.87 \%$ |
| \%C | $35.54 \%$ | $29.36 \%$ |
| \%NR | $23.69 \%$ | $23.82 \%$ |
| \%other | $1.93 \%$ | $1.03 \%$ |
| GPA | 2.685185 | 2.874317 |
| GPA II | 2.036517 | 2.182573 |


| All Students; MA 2051, |  |  |
| :--- | ---: | ---: |
|  | MA 2024 (n=588) | MA 1024 |
| \%A | $14.41 \%$ | $20.00 \%$ |
| \%B | $28.60 \%$ | $31.68 \%$ |
| \%C | $35.07 \%$ | $27.43 \%$ |
| \%NR | $20.46 \%$ | $19.47 \%$ |
| \%other | $1.46 \%$ | $1.42 \%$ |
| GPA | 2.735294 | 2.90604 |
| GPA II | 2.167373 | 2.332136 |

All Students less AP Math ( $n=456$ )
Math classes only

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $14.50 \%$ | $13.31 \%$ |
| \%B | $28.39 \%$ | $27.53 \%$ |
| \%C | $31.92 \%$ | $33.00 \%$ |
| \%NR | $17.59 \%$ | $20.13 \%$ |
| \%other | $7.60 \%$ | $6.04 \%$ |
| GPA | 2.767054 | 2.733305 |
| GPA II | 2.240306 | 2.147662 |

All Students; Math classes only ( $\mathrm{n}=588$ )

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $16.26 \%$ | $15.55 \%$ |
| \%B | $27.47 \%$ | $27.51 \%$ |
| \%C | $26.96 \%$ | $28.92 \%$ |
| \%NR | $14.62 \%$ | $17.46 \%$ |
| \%other | $14.69 \%$ | $10.57 \%$ |
| GPA | 2.84867 | 2.814224 |
| GPA II | 2.360588 | 2.264777 |


| All Math AP; MA 1024, $2051(n=132)$ |  |  |
| :--- | ---: | :---: |
|  | MA 2051 | MA 1024 |
| \%A | $20.18 \%$ | $35.25 \%$ |
| \%B | $35.09 \%$ | $47.54 \%$ |
| \%C | $34.21 \%$ | $9.84 \%$ |
| \%NR | $10.53 \%$ | $5.74 \%$ |
| \%other | $0.00 \%$ | $1.64 \%$ |
| GPA | 2.5625 | 3.274336 |
| GPA II | 2.54386 | 3.083333 |

AP Math, all classes, first year


AP math, all classes, four years


All Students less AP Math all classes，first year


All students less AP Math；all classes four years


$\square A$
$\square B$
$\square C$
$\square N R$
$\square$ other

All students; Math classes only, four years


## $\square A$ ■B $\square C$ $\square N R$ mother

AP Math; math classes only first year


回A
토․
$\square C$
$\square \mathrm{NR}$
$\square$ other

AP Math; Math classes only four years



All students less AP math, math classes only four years


| $I A$ |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |
| Dother |




All students; MA 2051


| 图A |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |
| $\square$ other |

* Other are classes students get credit for, but haven't taken.
** ( n ) represents the number of students in the category

| All Students; All classes ( $\mathbf{n}=\mathbf{5 8 8}$ ) |  |  |
| :--- | ---: | ---: |
|  | first year | four years |
| \%A | $23.65 \%$ | $32.12 \%$ |
| \%B | $28.92 \%$ | $29.89 \%$ |
| \%C | $22.10 \%$ | $19.35 \%$ |
| \%NR | $16.70 \%$ | $13.83 \%$ |
| \%other | $8.62 \%$ | $4.81 \%$ |
| GPA | 3.020715 | 3.156981 |
| GPA II | 2.468597 | 2.698451 |


| AP Chem; All classes | ( $\mathbf{n}=\mathbf{3 1}$ ) |  |
| :--- | ---: | ---: |
|  | first year | four years |
| \%A | $22.68 \%$ | $31.05 \%$ |
| \%B | $25.27 \%$ | $29.82 \%$ |
| \%C | $24.62 \%$ | $20.25 \%$ |
| \%NR | $16.20 \%$ | $13.20 \%$ |
| \%other | $11.23 \%$ | $5.69 \%$ |
| GPA | 2.973214 | 3.133174 |
| GPA II | 2.430657 | 2.694787 |


| All Students less AP chem <br> All classes <br> ( $\mathbf{n}=557$ ) <br> first year | four years |  |
| :--- | ---: | ---: |
|  | $23.71 \%$ | $32.18 \%$ |
| \%A | $29.12 \%$ | $29.90 \%$ |
| \%B | $21.96 \%$ | $19.30 \%$ |
| \%C | $16.73 \%$ | $13.86 \%$ |
| \%NR | $8.48 \%$ | $4.76 \%$ |
| \%other | 3.023358 | 3.158306 |
| GPA | 2.470708 | 2.698655 |


| All Students less AP chem; <br> chem classes only ( $n=557$ ) |  |  |
| :--- | ---: | ---: |
|  | first year | four years |
| \%A | $18.51 \%$ | $18.96 \%$ |
| \%B | $37.12 \%$ | $33.43 \%$ |
| \%C | $29.65 \%$ | $30.60 \%$ |
| \%NR | $11.48 \%$ | $15.11 \%$ |
| \%other | $3.23 \%$ | $1.90 \%$ |
| GPA | 2.869281 | 2.859659 |
| GPA II | 2.528802 | 2.419275 |


| AP Chem; | Chem clas | es only ( $\mathrm{n}=31$ ) | All stud | s; Chem | ss only ( $\mathrm{n}=5$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | first year | four years |  | first year | four years |
| \%A | 13.58\% | 20.13\% | \%A | 18.01\% | 19.04\% |
| \%B | 32.10\% | 32.70\% | \%B | 36.75\% | 33.38\% |
| \%C | 32.10\% | 29.11\% | \%C | 29.89\% | 30.48\% |
| \%NR | 8.64\% | 10.06\% | \%NR | 11.26\% | 14.74\% |
| \%other | 13.58\% | 8.18\% | \%other | 4.09\% | 2.35\% |
| GPA | 2.761905 | 2.892308 | GPA | 2.859734 | 2.861975 |
| GPA II | 2.485714 | 2.575342 | GPA II | 2.524013 | 2.429829 |



All less AP chem
All classes first year


| ■A |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |
| ■other |

All students less AP Chem
All classes four years

$\square A$
$\square B$
$\square C$
$\square N R$
⿴囗 other



AP Chem; Chem only, four years

$\square \mathrm{A}$
ロC QNR
$\square$
© other


| $A A$ |
| :--- |
| aB |
| OC |
| ONR |
| a other |

All students; chem classes only four years


| OA |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |
| Other |

## Computer Science

* Other are classes students get credit for, but haven't taken.
** ( $n$ ) represents the number of students in the category

| All Students; All classes ( $\mathrm{n}=588$ ) |  |  |
| :---: | :---: | :---: |
|  | first year | four years |
| \%A | 23.65\% | 32.12\% |
| \%B | 28.92\% | 29.89\% |
| \%C | 22.10\% | 19.35\% |
| \%NR | 16.70\% | 13.83\% |
| \%other | 8.62\% | 4.81\% |
| GPA | 3.020715 | 3.156981 |
| GPA II | 2.468597 | 2.698451 |


| All Students less AP Comp sci <br> all classes $(\boldsymbol{n}=573)$ <br>  <br>  <br> first year |  | four years |
| :--- | ---: | ---: |
| \%A | $23.40 \%$ | $31.75 \%$ |
| \%B | $29.19 \%$ | $30.11 \%$ |
| \%C | $22.20 \%$ | $19.49 \%$ |
| \%NR | $16.73 \%$ | $13.90 \%$ |
| \%other | $8.48 \%$ | $4.76 \%$ |
| GPA | 3.016037 | 3.150742 |
| GPA II | 2.464592 | 2.690778 |


| AP Computer Science; <br> All classes $(\mathbf{n}=15)$ <br>  <br> first year |  |  |
| :--- | ---: | ---: |
| four years |  |  |
| \%A | $32.53 \%$ | $46.16 \%$ |
| \%B | $19.68 \%$ | $21.98 \%$ |
| \%C | $18.88 \%$ | $14.30 \%$ |
| \%NR | $15.66 \%$ | $10.92 \%$ |
| \%other | $13.25 \%$ | $6.63 \%$ |
| GPA | 3.19209 | 3.386435 |
| GPA II | 2.615741 | 2.990251 |

All students; CS class only ( $\mathrm{n}=588$ )

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $19.51 \%$ | $27.37 \%$ |
| \%B | $30.82 \%$ | $30.28 \%$ |
| \%C | $21.29 \%$ | $23.34 \%$ |
| \%NR | $21.95 \%$ | $17.18 \%$ |
| \%other | $6.43 \%$ | $1.84 \%$ |
| GPA | 2.975232 | 3.04973 |
| GPA II | 2.277251 | 2.516065 |

AP Computer Science;
CS classes only ( $\mathrm{n}=15$ )

|  | first year | four years |
| :--- | ---: | ---: |
| \%A | $41.94 \%$ | $59.41 \%$ |
| \%B | $16.13 \%$ | $18.81 \%$ |
| \%C | $9.68 \%$ | $6.93 \%$ |
| \%NR | $12.90 \%$ | $7.92 \%$ |
| \%other | $19.35 \%$ | $6.93 \%$ |
| GPA | 3.47619 | 3.616279 |
| GPA II | 2.92 | 3.308511 |

AP CS all classes; first year

-A

- B

ロC

- NR
- other

AP CS all classes; four years


All students less AP CS; all classes first year


图A

All students less AP CS; all classes four years


All students; cs classes only first year


ロC - NR E Other

All students; CS classes only four years


## All students less AP CS；CS classes only first yea



All students less AP CS；CS classes only，four years


| 图A |
| :--- |
| ■B |
| $\square \mathrm{C}$ |
| $\square N R$ |
| $\square$ other |

AP CS; CS classes only first year

$\square A$
$\square B$
$\square C$
$\square N R$
I other

AP CS; CS classes only four years


| 四A |
| :--- |
| $\square B$ |
| $\square C$ |
| $\square N R$ |
| ■other |


[^0]:    ${ }^{1}$ Binomial trials-consists of $n$ independent success or fail trials. And the probability for success is the same value, p , for all trials.

[^1]:    ${ }^{1}$ Normal Distribution-a distribution function which follows the bell shape curve. Provides a method for approximating other distributions, like binomial.

[^2]:    ${ }^{1}$ This p-value is too high to say that the AP students have a significantly different NR percentage. The percentage we would've needed to be able to say the difference was significant would be around $95 \%$, yielding a p-value of .05 .

[^3]:    The p-value of about .2 doesn't mean our results are useless, it just means the difference in the NR percentages is not statistically significant. This means that there is not enough evidence to disprove $\mathrm{H}_{0}$.

