Packet P3 -

Decide on a Project and Write the Introduction

Over the next 10 days, you will choose a general topic or theme, narrow down it down to a specific focused task, and write your introduction. Consider the following:

- ✓ Start by brainstorming several general areas of interest or curiosities in your life. Eventually, you will pick one and then narrow it down to a specific task or focus.

- ✓ No matter what you eventually end up with, you must be able to collect or generate raw data related to your idea. There are two option, primary data, secondary data, or both.

- ✓ Please note that this is not a science project where you make a hypothesis. You may have a hypothesis if you want or you can just say that you are investigating something.

- ✓ Read the packet P2 (Posted Online) on ***Generating Data for your Project*** *which discusses some factors for you to consider.* You will eventually have lists of raw data that you will submit as part of your project as well as send it electronically to the project adviser.

  For Primary data:

  - Think creatively about possible surveys or experiments you could set up which support your investigation.

  - All surveys need to be pre-approved by Mr. Cedarlund before they are handed out. To do this, include the survey when you turn in your introduction. You may want to drop by to discuss it ahead of time.

  For secondary data (data from trusted sources, gov't, academics, industry, etc)

  - You must be able to access raw data/information that is not already summarized into tables or graphs or statistics. Hopefully it can be easily transferred into a spreadsheet for huge time savings and flexibility benefits. This may be possible by copying and pasting into a spreadsheet program like Excel or Google Sheets.

  - Research the type of data that is available for your topic completely before you decide the **focus** of your project.

  - Search, thoroughly, the related literature on the web, library, the media, and other resources. Don't just assume you can get certain data on the web. Go actually find some actual sources to be sure you can clearly find what you will be using. Remember, you need raw data and not data already summarized into graphs.

- If using secondary sources, you are required to provide a bibliography and footnotes. Write down the *www* address(s) as a record.

Keep track of all decisions you make as you track and narrow down your data. You will eventually, as part of your project, have to discuss your decisions you have made in narrowing down your final data.

General Advice as You Search for ideas

- ***Choose a project that you truly find interesting*. Do your own work. Make this your own project.**

- Always keep the Project Assessment Criteria in mind. (criteria A to G). To get high scores (5, 6, or 7 *out of 7*), this entire project should be a major piece of work and should demonstrate commitment of time and effort.

- Projects should include the bare minimum of 2 *simple* math processes, but it is a good idea to include additional *simple* math processes.

- Higher scoring projects will require at least one "further" math process. You can include more. For at least one further process, you will have to demonstrate a highly detailed and in depth understanding of the process.

- Additional "simple" or "further" processes can be included to possibly help deepen the level of analysis and sophistication of the project, possibly raising your score. Every single feature or process you include must be relevant in a meaningful way to the main task of your project.

- Don't include any process, mathematical or otherwise, unless it is fully relevant. When deciding whether to use a math technique: Ask yourself *if it is genuinely useful in order to investigate my task* in a meaningful way.

- **The highest scoring projects:** Just having lots of math processes alone will not guarantee you upper level scores on this project. For those scores, you will need opportunities for robust analysis. This can be enhanced by including additional variables of data and setting up your project so that you will be forced to be thoughtful and comprehensive in your analysis. These projects take a lot of time, especially if you desire a score higher than a 4.

- Consider keeping a diary of your work, making a note of your thoughts and ideas for later inclusion in your write-up. Keep track of any prospective weaknesses that you notice in your project and weaknesses in the mathematical choices you are making. (this will help you to question your results later and earn points in the Validity criteria)

- Lastly, collaboration with peers/tutors/teachers is allowed only to receive guidance and assistance. Plagiarism is taken seriously. *Turnitin.com* among other resources will be used to monitor projects.

Potential Mathematical Skills to Utilize

Simple Math Processes (include 2 or more) :

- Percentages, Rates, Proportions *(one could also take raw data and calculate a rate or a %. That value could then be used as a variable in your project to do something else),* Comparing ratios or percentages that you calculate.
- Areas or volumes of plane shapes
- % error (nice for comparing)
- Right triangle trig, Non right triangle trig -- Law of Sines, Law of Cosines
- Pie charts, Bar charts, Dot Plots, Stem plots, etc
- Scatter plots to estimate correlation visually or for some other purpose *(need at least 30 to 40 points minimum on a single scatter plot)*
- Scatter plots to help you to create a mathematical model (equation) in order to make predictions.
- Frequency tables *(won't count as a standalone process but can help lead to histograms or cumulative frequency graphs, etc.)*
- Histograms
- Cumulative frequency curves ("S" curves)
- Mean, Median, Mode *(each of these should be a separate process each is used in a relevant way). Don't lump them together)*
- Range and IQR
- Standard Deviation (done in conjunction with the mean to help compare variations in data).
- 5 number summary / box & whisker plots (based on the median)
- Percentiles
- Probability
- Arithmetic & Geometric Sequences

Further Math Processes (try to include at least one) :

Must show an in-depth understanding of the process in terms of your information including in-depth, step by step, calculations with explanation. Note: Any of the further processes can also count as simple math processes if you need additional simple math processes.

- Use the Chi Squared Test of Independence (you really need a lot of data to be safe)
- Pearson's Product Moment Correlation coefficient, r.
- Least Squares Regression Line (LSRL) Note: don't use the term "Line of Best Fit" (The LSRL could be relevant if you are going to use the equation to make predictions if the trend is linear AND if there is at least moderate correlation. One could also use the slope of this line to analyze the rate of change of a linear situation. Don't include the LSRL process just so you can place a line on top of a scatter plot with no other purpose).
- Create a mathematical model from the data - create and use functions from real life data (linear, quadratic, exponential, etc)
- Use *residuals* in conjunction with a Modeling project
- Analyze exponential functions of data that appears to be exponential.
- Calculate Probability using the Normal Distribution.
- Analyze data that is periodic and create Periodic functions in the form as $y=A\sin Bx + D$, etc
- Use Compound probability

If you use Correlation or LSRL in a project:

Correlation can indicate the ***degree of linear strength*** between two numerical variables. Correlation can be viewed visually on a scatter plot and more specifically with *r*, Pearson's Linear Correlation Coefficient. Remember that correlation does NOT imply causation.
If you draw a scatter plot and there is no correlation, it is still allowable (and relevant) to calculate the correlation coefficient, r, to verify the fact. However, it would not be relevant in that case to calculate the LSRL (regression line). If the correlation is strong enough, then you may calculate the regression line equation and use it to count as one of your sophisticated processes assuming your LSRL is relevant and part of a useful plan to begin with.

If you do not draw a scatter plot, the relevancy of a regression line will depend on the value of *r*, the Pearson's Product Moment Correlation Coefficient. If you simply just write down the value of *r* from your calculator, then this will be graded as a simple math process only. The LSRL could count as a further process if full details are shown.

A linear regression equation (LSRL) would ONLY be relevant if there is evidence that there is at least a moderate correlation and you are intending to use the equation for modeling, extrapolating, and conveying information. Evidence can be from a scatter plot or from the correlation coefficient. Don't include a LSRL if you are not going to make use of it.

If you use the The Chi-Square Test of Independence :

The test is not a study of correlation. The Chi Square test does not prove anything. It can lend evidence to support two categorical variables being associated with each other. If carefully done and with common sense, numerical variables can be converted to categorical variables. But if both variables are quantitative, then you should probably not use the Chi Square test. Note: Even though the Chi-Square test indicates whether two variables are associated, the test cannot indicate how strong the association is.

NOTE:

IB advises against calculating a correlation coefficient coefficient AND conducting the Chi-Square test of independence with the exact same set of 2-variable data. If you do one, the 2nd procedure may be deemed as not being relevant and lower your score. You could do both math processes as long as a different combination of data is used.

If you create a Mathematical Model from your data:

Scatter plots would most likely be involved as a starting process to view and analyze the pattern. This could lead to the use of many types of mathematical functions, not just linear (quadratic, exponential, etc).

To count as a further process, there would have to be discussion/analysis done to determine the equation. If just a regression button on a graphing calculator or Excel is used, then the process would only be counted as simple.

Write your Introduction and define your variables

Your introduction and variable definition section you are about to write is essentially the one that will part of your final draft (after some feedback and final editing). Follow the guidelines below **and** the first part of the "***In-Thinking" Project Writing Guide.(pages 1-4).***

Carefully think ahead on how you will carry out this project **before** you write your introduction. Your introduction clearly tells the reader what will happen in the rest of your project. Edits to the introduction can be made in later stages of the project. There is a soft 2000 word limit on whole entire project. Your project, starting with the introduction, should read well, be concise, be logical, and should not contain "fluff".

Example of Correct Header (does not need to be on a separate Title Page until Draft #2)

Jonas Peterson IB # 000440-231
Eugene IHS
IB Math Studies SL
Exam Date - May 2019
Title: What Factors Contribute to Winning in the WNBA ?

Draft #1 – Introductions and Definitions

Table of Contents (not necessary yet unless you want to get the beginning of it set up now)

Paragraph 1 Statement of Task (The What)

Describe the overall purpose of your project and what you hope to achieve. You can include brief reasons why think it is interesting if you want. Write in future tense.

Include enough detail so that the reader has a good handle on the overall scope of what and who is involved. Don't bombard the reader with a mass of mathematical terms in this first paragraph. *(Save that for the Plan). Compare the three statements below.*

A. *I will investigate if GDP is independent from number of deaths.*

B. *I intend to investigate if the GDP of a country had a possible affect on the number of deaths that occurred.*

C. *I want to see if the richest countries during WWII were able to possibly hire more people or possibly generate more weapons to assist with their murders.*

Notice statement A is very technical and not too descriptive. Statement B is decent and more reader friendly. Statement C clearly conveys the purpose and is in easy to understand, every day language.

Be Careful of your Project Title - It can really limit you!
Read pages 1-4 in the writing guide about Project Titles and projects that are too limiting.

There is a bit of a paradox when your title is "Is there a correlation between _____ and _____". A title like this is very limiting as any other math process other than correlation is irrelevant which means you eliminated 95% of the mathematical and statistical math processes you have available.

Paragraph 2 About your data

Briefly mention where your data will come from. If secondary include the official name of the organization and its website.

I plan to collect the cancer rates from the the CDC, Center for Disease and Control….

Or … I plan to create and distribute a survey to random people at the shopping center in order to collect….

Clearly specify exactly what data/information you will be collecting and be very specific.

I will collect the number of people with bone cancer in each of the countries mentioned above. I will also collect the population of each state from…..

Or…I will collect number of points scored over the 2015 and 2016 regular season from a random selection of players in the league. I will also record the position they play. I will also collect their salaries over this same period.

Note about data Quality
The quality of data in some projects can suffer when you limit the scope of your data collection. *For example, someone could study the association between the weight of American football players and how many tackles a player makes. If you only look at the top 50 tacklers (or only the 50 heaviest players) because it's convenient, you might be overlooking some patterns.* Unless the project truly needed to focus only the heaviest players for some reason, the quality of the data suffers a bit. This could potentially drop your project score one or two points. *This person would be better off to sample a broader range of weights (or # tackles) to better represent the data. They could get a random sample of all players or take a stratified sample by taking every 5th player for example.*

Paragraph 3 The Plan (The meat and potatoes portion of the Introduction)

This is the "**HOW**" part of the introduction. Lay out a detailed **plan** of what you intend or anticipate doing with your information in an order that makes logical sense.

Use a numbered (or bullet point) list for each math process. Order is important. For each bullet point or number, include a single math process you will do AND explain the rationale for using the process. You do not need to be highly detailed about each technique but you do need to name the math process and explain its purpose or relevance as it relates your task. If appropriate, specify the portion of your data that you would use for that math process.

The following are sample statements from a variety of projects:

- *I will use the median sprint times over the season to create box plot for each of the three teams. I will use the box plots to see which teams tend to have the fastest and most consistent times.*

- *I will calculate the mean average heights of 12 year old boys and compare it to heights of 13, 14, and 15 year old boys to see the rate at which they are getting taller. I will then repeat for the females.*

- *I will create a scatter plot to visually assess the correlation between the weight of my rocks against the cost of the shipments. The weight of the rocks will be the independent variable.*

- *I will perform the Chi-Square Test of Independence to determine if the type of car and the type of bumper sticker are associated.*

- *I will calculate the mean and standard deviation of SAT scores. Assuming a normal distribution, this will allow me to calculate the probability of one of my participants in my survey of having an SAT score of more than 1200.*

Paragraph 4: Definition Section

This is not actually a paragraph but a list of definitions of your variables. For every column in your data table that you will show in your project, you should have a clear definition in this section. Define these *numbers* in terms of *your* project, don't just write the dictionary definition. If a number has a unit be sure to include it.

Examples:

Crime rate: *The number of crimes in a state for every 1000 people in the state. Crimes include…..*

Throwing distance: *The distance that a ball was thrown by the participants in my study, in feet and inches. Later in the project I will covert to a singular unit of inches.*

Happiness Index: *A value from 1 to 8 as determined by the happiness scale I have created. See the survey.*

Other requirements and Tidbits:

- Be sure to use appropriate mathematical terminology in your introduction.

- You must *word process* your introduction.

- **Surveys:** If you are creating a survey to be distributed, then you should create a well thought out rough draft to submit to me with your introduction. Don't pass out you're your survey or collect data until you get my feedback. Think ahead of what steps you can take to get a reasonable random sample of unbiased data.

The test of a good introduction: I could give your intro to anyone in the class and they would know what to do without asking you any questions. Introductions submitted that have been written quickly with minimal thought will get you off to a bad start to your project.

Your class grade for "Draft 1 - Introduction and Definitions"
After reading your Draft 1, Mr. Cedarlund will assign it a class grade **(out of 60 points)** that will apply during your trimester 1 grade in the Test/Project Category. The grade will be based on three things:

a. Your project has been thoroughly planned and well thought out.
b. You followed the details of this Packet, P3 and the Writing Guide (pages 1-4)
c. You meet deadlines.

This grade does not count toward your "IB" Grade toward your project. Your IB grade will come exclusively from your Final Draft of your entire project.

Turn-In "Draft 1 - Introduction and Definitions"

by **Monday, October 14th**. You will submit both a copy via Turn-it.com and submit a paper copy. Turn-It in Code is _____

- I can only give a limited number of projects my attention each night. Therefore, I provide feedback in the order I receive them, with seniors getting top priority. The earlier you turn it in, the quicker I can return it to you and the more time you will have to do the next phase of the project. We have a very large class so don't expect immediate turnarounds.

- Seniors. I realize that you have a large IB deadline close to the time Draft #2 is due. If I were you, I would turn in your introduction on the early side.

- If doing a survey, be sure to include a well thought out survey you are thinking of using.

Length

The approximate maximum word count of the final draft is 2,000 words, *excluding diagrams, graphs, appendices, and bibliography.* However, it is the *quality* of the mathematics and the processes used and described that is most important, rather than the number of words written.