

Format:

We will follow the APA style of formatting. In the APA style of formatting, you should use **standard-sized white paper, (8-1/2 X 11) with one-inch margins at the top, bottom, right and left sides and a double spacing format**. For more info on APS style, take a look at any of the following sites:

<http://www.ccc.commnet.edu/apa/format.htm>

<http://www.docstyles.com/library/apalite.pdf>

<http://www.ccc.commnet.edu/apa/format.htm>

The project should follow the detailed guidelines that I have stated earlier. For each section, please read the instruction and just follow the guidelines. Please note that the project is divided in two separate but related parts. The first part is the Analysis Part and second part is the use of Excel in generating charts, tables and numbers. The second part of the project will be reported in the appendix section of the computer project. The first part (Analysis Part) focuses on getting to know the topic, identifying the purpose, explaining the background, stating the limitations and or policy implication of the topic; and more importantly, the verbal descriptions of the charts, numbers and tables. You should not include any charts or tables on the Analysis Part. Again, the second part includes all the charts, tables and numbers that are created by Excel. Please note that each table or chart is already numbered and your typical analysis starts by “As Table 2 indicates” Or, “Chart 5 suggests that”

The analysis part is the heart of the report. It contains 10 sections i.e. the Topic Section, the *Background Description* section, the *Scope of the Project* section etc. For each section, read the detailed guideline and report what is needed for that section of the report.

Guidelines (sections etc.):

Computer projects should be presented in a *professional format*. **Let's focus on the most recent measurement of your dependent variable**. Repeat: on project #1, you are going to apply tools and techniques of descriptive statistics ONLY on your **dependent variable**. Other variables will be used in projects # 2 and #3.

You will use other variables in projects 2 and 3.

Let's use the following format and include the following sections to present your analysis. Your report should encompass the following sections:

1- A professional looking **cover page** with proper graphics to represent the subject.

2- Table of Contents

3-Topic - Describe the topic and variables selected. Why are you interested in this topic? In another word, describe the importance of the topic and explain why you have selected this topic. Relevance, practical applications and ingenuity are important considerations. You should explain its relevance and discuss its practical applications. What are the policy implications of your hypothesis?

3-Background Description - What is the history of the topic? Provide a detailed description of its historical background along with relevant historical information i.e. *When, How, What, for Whom etc.* Use Internet and conduct a short **review of literature** regarding your topic i.e. find out if there are other studies that are “similar” to your research. **Include a brief description of their findings.** Provide the proper citations for your review of literature. (At least 3 pages).

3- Statement of the Research Objective, i.e., The Scope of the Project –What is the main objective of your project? What is the secondary objective of your project? Statement clarity is very important. A common way to begin the statement would be to say, “**The main purpose of this project is to explore, analyze, and describeVar1.**

The secondary objective of this project is to examine possible relationship between Var1 and VAR 2”.

(This part is only a short paragraph)

5- Data Source –

Specify the source of your data. Where, when etc. Explain if your data is a time-series or a cross-sectional data. (Again, this part is only a short paragraph)

6- Statistical Analysis, i.e., Application of the Descriptive Methods –

This section is the heart of your computer project.

Although you have collected measurements on a time-series base (for two time periods), focus **only** on the most recent measurement for statistical analysis. Select relevant descriptive methods that can help you to achieve your **objective** and apply proper statistical methods that are optimum for your variable. Tabular, graphical and numerical methods **should all be utilized to describe your data set.**

Your job is to **explain and verbally interpret every table, every chart and every number generated by Microsoft Excel in details.**

The grading rule is simple: more verbal interpretations, the higher the score. **Again, the rule of thumb is: the more tools and techniques you use, the more points you will get.** So, apply every tool, technique and method that you have learned in chapters 1-3 to explore, analyze, and “describe” your variable.

Computer generated tables, graphs and numbers should be presented in an appendix. If needed, Exploratory Data Analysis and the cross-tabulations and scatter diagram should be used to explore the relationship between variables. Make sure that computer printouts and graphs are placed at the *end* of the report.

Tables and graphs should be presented in a professional manner i.e. a descriptive topic for each table or chart, 3-D graphs with proper legends etc. By assigning a number to each graph and table, which will appear in the appendix section, you can refer to these graphs and tables to state” As table # -- indicates.....” or “Graph #... shows that..... “ then go ahead and describe what that table or graph indicates.

The organization of your analysis should follow the following format (*same heading for each section*):

6. Statistical Analysis

I. Numerical Methods:

a. Numerical Measures of location, dispersion, and shape of distribution.

Numerical measures such as descriptive measures of location and dispersion for the most recent measurement should include the interpretation of: the Mean, the Mode, the Median, the Standard Deviation, the Max, the Min, Q1, Q3, the Range, the IQR, the Coefficient of Variation, the Z-score, the Percentile (choose 20th, 33rd, 66th and 90th percentile), the Skewness, the Kurtosis, and the application of the Chebyshev's theorem. Also, as a practice, use the Empirical Rule, even if the distribution of your data set would not be the required Normal distribution.

Make sure that your interpretations are clear, complete and in the context of your topic.

b. Numerical Measures of Association

Use correlation analysis to comment on the possible relationship between your Dependent Variable (D.V.) and the Independent Variable(s) (I.V.)

The Correlation Coefficient (r) is a measure of the strength as well as the direction of a linear model to observed data. The value of the (r) is always expressed as a percentage and it varies between $-1.00 \leq r \leq 1.00$

Overall, the larger the r , the stronger the relationship

Tips regarding the DIRECTION of the linear relationship:

If the sign of the “ r ” is positive, then use “Direct”, and if the sign of the “ r ” is negative use “Indirect” as an adjective to describe the relationship.

Tips regarding the STRENGTH of linear relationship:

Disregard the sign, if the absolute value of the “ r ” is within the following ranges, then use the suggested adjectives to interpret your findings:

$r > 0.70$ Strong, $r > 0.85$ Extremely Strong;

$r < 0.30$ Weak, $r < 0.15$ Extremely Weak, $0.40 < r < 0.60$ Medium,

$0.60 < r < 0.70$ Semi-strong; $0.30 < r < 0.40$ Semi-weak

II. Graphical Methods:

a. Graphical Measures

Graphical measures for the most recent measurement should include the following professional looking graphs, as well as complete interpretation of:

The Bar Graph, the Pie Chart, the Dot Chart, the Stem and Leaf Display, the Box-Whisker Plot, the Polygon (the area chart), and the Ogive (use the ogive to estimate percentage that are at least and at most a specific value that you choose).

All the tables and graphs should be developed on a professional base.

Make sure that your interpretations are clear, complete and in the context of your topic.

b. Graphical Measures of Relationship – Visualization Step

Develop the scatter diagram for the data set. Comment on the possible relationship between X and Y. The use of the scatter Plot allows us to make a comment on the possible relationship between X and Y:

X and Y appear to be:

- ✓ Not related
- ✓ Related in a non-linear fashion (Polynomial, Exponential, Log etc.)
- ✓ Related in a linear fashion :
 - I. Is it a direct linear relationship?
 - II. Is it an indirect linear relationship?

Do you see any pattern? What does the pattern look like? Does it appear to be a linear or a non-linear relationship? If linear, does it appear to be “direct” or “indirect” linear relationship?

IV. Tabular Methods:

a- Frequency , Relative Frequency, Cumulative Frequency and the Relative Cumulative Frequency Distribution Tables.

Tabular methods for the most recent measurement should include the interpretation of: the Frequency Distribution Table, the Relative Frequency Distribution Table, the Cumulative Frequency Distribution Table and the Relative Cumulative Frequency Distribution Table.

Let's imagine that you are a reporter. Avid over or under-analysis of tables. Use the guidelines that I provided in class. Wherever possible, convert percentages into ratios. For each table, the following comments are necessary:

- a- Make a comment about any newsworthy finding for each table.
- b- Compare and contrast different categories and make some comments.
- c- Aggregate couple of categories and make a comment.
- d- Compare and contrast aggregated categories.
- e- Use ratios where applicable.

b- Tabular Measures of Relationship – Cross tabulation

Analysis of relationship - Use joint frequency distribution or “cross tabs” to comment on the possible relationship between your Dependent Variable (D.V.) and the Independent Variable(s) (I.V.). It would be a good idea to express the percentages and the “Total Percent”, the “Row Percent”, the “Column Percent”.

Make sure that your interpretations are clear, complete and in the context of your topic.

7- Exploration of Possible Changes.

Interpretation of the changes in the summary statistics (Numerical Methods) - the 13 numbers (concepts)- Mean, Mode, Median, St. Dev. etc. should be interpreted in details.

For each **concept**, there are 3 possibilities:

- An Increase (expansion, improvement)
- A Decrease (deteriorated, worsened, depreciated, weakened)
- No change (Stayed the same)

Note: The idea of this section is not simply repeating the interpretation of the concepts in 2 different time periods. The analysis of change requires comparison of such numbers and analytical comment on possible change (variation) that has occurred on each concept.

In another words, the idea of the “analysis of possible change” has two parts:

- 1- You have to know the CONCEPT that you are trying to analyze. What does the CONCEPT measures?

The concept (Mean, Mode, IQR etc.) that you are analyzing is designed to measure certain statistics. For example:

The Mean measure the expected value for a typical element, the IQR measures the spread of the middle 50%, The Kurtosis measure the degree of increase or decrease in the peakness (flatness) of a distribution etc.)

- 2- What was the magnitude (the amount) of change in the concept being measured? The direction and the magnitude of the change in the concept being measured. Did the concept (Mean, IQR, etc.) show:
 - a- An increase (improvement, progress, advancement, enhancement etc.)
 - b- A decrease (deterioration, decline, depreciation, worsening, fall, drop, etc.)
 - c- No change – The concept being measured stayed the same (unaffected, constant, alike, etc.)

8- Discussion of the Possible Outliers - Explore and describe mild and or extreme outliers that may appear in your data. If your data is normally distributed, you can use the Z-Score to evaluate the existence of possible outliers. Otherwise, develop a box-whisker plot (that will be included in the appendix section) to answer this part. Determine the outer and the inner fences and identify the “mild” or the “extreme” outliers. Identify the smallest and the largest values within the “inner” fences.

9. Limitations of the Study - State limitations that may have affected the analysis, i.e., non-current data, difficulty of measuring the concept in question etc. In part 2 you have explained your model in terms of the “cause and effect” i.e. one dependent (or response) and one independent (or explanatory) variable in your simple model. In a multi-variable model, you could have included more than one independent variable to explain variations in your response variable. State the list of other relevant explanatory variables that could have enriched your model.

- 1- **Conclusion, Summary, and Policy Recommendations** - This is where you tie everything together. Provide an executive summary and interpret your results. Comment and include all newsworthy (timely and important) information.

12- Appendix –

Please note that every Excel generated graph, table, numbers, and charts should be reported in the appendix. This section should include your computer printout (tabular, graphical and numerical). Include the bibliography for the review of literature of the “**Background Description**”

There is no need to show the actual data.