

## Assignment: Data Analytics Case (Written)

wayne.smith@csun.edu

[ updated: Wednesday, March 27, 2024 ]

---

**Course:** BUS 312

**Title:** Data Literacy for Business (3 units)

“The best way to predict the future is to create it.”

-Peter Drucker (1909-2005)

### Goal

The purposes of this assignment are as follows:

- To analyze data related to a dataset with breadth and depth
- To identify one or more important business issues exhibited in the data
- To draw out ethical issues of importance related to this data
- To use skills, knowledge, abilities from this course and prior courses
- To make actionable recommendations to leverage opportunities and overcome obstacles
- To continuously improve the practice of writings in substantive business contexts

### Instructions

In general, your team’s task is to analyze a dataset, and the ethical issues that might arise in the dataset or the related business contexts.. Essentially, your team is analyzing key data issues.

Both team composition and selection of a dataset/question will be done before the middle of the semester. Each person in each team must contribute both to the analytics and to the writing. The written deliverable must be printed and delivered to the instructor at the beginning of class on the due date. The written deliverable must also be submitted electronically in Canvas no later than 1 hour before the beginning of class on that day.

### Deliverable

The due date is listed on the course outline. Before that due date, there should be enough time in or out of class to go over any questions your team may have.

### Length

The “Main Body” of the written deliverable must be no less than ten (10) pages and no more than twenty (20) pages (not including the cover page, cover letter, executive summary, references, or appendices). The written deliverable may include an Appendix or may not include an Appendix depending on circumstances.

## **Formatting**

Use only double space. You can choose to indent paragraphs if you like. Use 1 inch margins (top, bottom, left, and right).

## **Scoring**

There will be 100 points allocated for breadth and depth of content. There will be 100 points allocated for language use and grammar.

## **Context, Resources, and Audience**

The following is excerpted and adapted from Professor Roger Peng's writings at:

<https://leanpub.com/dataanalysisessays>

The three key aspects of data analysis are context, resources, and audience.

The context of a problem covers many factors, including the question that gave rise to the dataset and the motivations of the people asking the questions. Understanding what kind of solution needs to be created and what kind of answer might be useful is a key part of the context. Datasets and statistical methods can be deemed useful or not based solely on the context of a problem. Data analysts must make an effort to understand the context of a problem in order to develop a complete understanding of the problem. Getting a better understanding of the context may require communicating with subject matter experts or other people involved in posing the question.

Every data analysis is conducted with a limited set of resources. The resources available determine the completeness and complexity of the solution developed. The most critical resource for a data analyst is time, but computing power, technology, and money also play a critical role. Understanding what resources were available when a data analysis was done allows us to evaluate the quality of an analysis relative to what else could have been done.

Who is this data analysis for? Every data analysis has an audience, even if it is the data analyst alone. Knowing who the audience is and what they need is important for determining whether an analysis is successful or not. Further, if one considers data analysis has a rhetorical activity, then it can be important to understand how an audience receives information and evidence. Many factors can play a role here, including cultural differences, differences in background knowledge, and technical proficiency. While it may not be possible for a data analyst to know all of these things as an analysis is being conducted, the analyst must make a reasonable attempt to obtain this information.

## **Content**

In general, the structure of the content ("main body") of the written deliverable is to follow the general outline and sequence of the scoring rubric. Use your own words, and not textbook

chapter titles. Apply the concepts from the textbook chapters and other materials from class that help your team analyze the data and address the business question. In addition, of course, to material from this course, this assignment builds upon ideas from several general education, lower-division core, and upper division courses.

The following document contains additional details regarding this written analysis including the correct sequence:

<https://ocw.smithw.org/bus312/casewritingrequirements.pdf>

Use a “References” (or “Works Cited”) page. Make sure your team also uses “In-Text” Citations as well. The specifics of the citations are enumerated in the document above.

## **Introduction**

Provide an introduction and overview of each member of the team. What skills, knowledge, and abilities does each team member bring to the team? How do those skills, knowledge, and ability contribute to a business decision or use data? Which skills, knowledge, or abilities are from education and which are from experience?

Provide an introduction and overview of the case choice. Why was it selected? What were the other topics considered? What was interesting about those other topics not selected?

Summarize two non-class articles broadly related to the topic. The articles must come from one of the sources listed at the end of this document.

Provide an introduction and overview of the dataset. Summarize the data including its relevance and source.

## **Descriptive-level Analytics**

An analytical workflow begins elementary analysis that simply describes the data in various ways—usually summary tables and graphs (charts). Descriptive Analytics doesn’t answer questions; Descriptive Analytics often raises questions or revises initial questions.

## **Visualization**

Your analysis must include at least four graphs. At least one of the graphs must be a histogram. The graphs must be labeled correctly, including appropriate x- and y-axis labels, and titles. Practice making sure the graphs are situated correctly in the written and can be seen well in the presentation.

## **Exploratory Analysis**

Your analysis team must provide two different exploratory analyses. For example, one exploratory analysis can be on one field (column) and another exploratory analysis can be on another field. Or, each exploratory analysis can be on two different datasets. Follow the process we used in class, and learned in your elementary business statistics class, to summarize the central tendency (e.g., mean or median) and dispersion (e.g., range or standard deviation) of a field (this is sometimes called “univariate” analysis). You must also provide at least one frequency distribution (i.e., a table of counts or cross-tabulations). A frequency distribution enables a summary view of the entire field as classified by certain categories.

### **Preliminary, Interesting Information**

What preliminary information did the team discern from the visualization and exploratory analysis? What questions does it raise? Which ones are the most important?

### **Diagnostic-level Analytics**

An analytical workflow continues with digging deeper into one or more questions. In particular, we need to use hypothesis testing to “rule out” randomness. That is, we want to have some confidence that data we see is more than just being random.

### **Decision, Question, or Issue**

What decision, question, or issue do you want to analyze? Of all of the possible questions that your team might want to ask, why is this one the most important?

Justify your question. Your team must use at least one idea from the four upper-division core courses (i.e., FIN 303, MGT 360, MKT 304, or SOM 306); or, at least one idea from five of the six lower-division core courses (i.e., Principles of Financial Accounting, Principles of Managerial Accounting, Principles of Micro-Economics, Principles of Macro-Economics, Business Law—we’ll exclude Business Statistics). See Sources for LDC/UDC concepts (below). In addition, your team members’ own textbooks and course materials, and the textbooks on reserve in the Oviatt Library Reserve Bookroom may be helpful.

### **Explanatory Analysis**

Perform and explain a complete Hypothesis Test (e.g., a two-sample t-test). Perform and explain a complete Confidence Interval. In general, a hypothesis test and a confidence interval support each other. This confirms, or explains, your data in the context of the question.

### **Compelling Finding**

What compelling finding can you conclude from the results of the Hypothesis Test and Confidence Interval?

### **Predictive-level Analytics**

An analytical workflow continues even further with digging deeper into one or more questions. In particular, we would like to forecast some new information. This is done by using the existing information. You will use a regression of some type.

### **Decision, Question, or Issue**

What decision, question, or issue do you want to analyze? The question is going to be different from the diagnostic analysis but it can come from the same or similar dataset. Of all of the possible questions that your team might want to ask, why is this one the most important? Since this is different part of the analysis, and a different question, it requires a different kind of analysis.

Justify your question. Your team must use at least one idea from the four upper-division core courses (i.e., FIN 303, MGT 360, MKT 304, or SOM 306); or, at least one idea from five of the six lower-division core courses (i.e., Principles of Financial Accounting, Principles of Managerial Accounting, Principles of Micro-Economics, Principles of Macro-Economics, Business Law—we'll exclude Business Statistics). See Sources for LDC/UDC concepts (below). In addition, your team members' own textbooks and course materials, and the textbooks on reserve in the Oviatt Library Reserve Bookroom may be helpful.

### **Forecast Analysis**

Perform a Correlation Analysis (this is sometimes called "bivariate" analysis). Additionally, perform a Regression Analysis. You can use either simple (single x-value) linear regression or multiple (more than one x-value) linear regression. Remember to clearly understand the slope and intercept, and R-squared, well.

### **Compelling Finding**

What compelling finding can you conclude from the results of the Correlation Analysis and the Regression Analysis?

### **Ethical Analysis**

Discuss at least one ethical issue that arises in your team's analysis of your selected dataset.

### **ESG or U.N. Global Compact Classification**

Ethical issues can be broad and expansive but to analyze them, much less make actionable recommendations, need to be clear and explicit. Classify the ethical issue or issues. 1), Your team can use one or more of the four U.N. Global Compact categories (Human Rights, Labor, Environment, or Anti-Corruption) that we discussed briefly in class. Which category does your issue or issues fall into? Be specific. Which principle in which category does your team feel applies the most?

<https://www.unglobalcompact.org/what-is-gc/mission/principles>

Or, 2), Your team can use one or more of the three Environmental, Social, or Governance (ESG) categories that we discussed briefly in class. Which category does your issue or issues fall into? Again, be specific. Which specific element in which category does your team feel applies the most?

<https://research.ftserussell.com/products/downloads/ESG-ratings-overview.pdf>

## Normative Ethics

CSR theory, in turn, is based up Normative Ethics. Normative means “should”. Your team must use at least one ethical theory (Utilitarianism, Deontology, Nicomachean, Rights, or Justice, but not CSR because that was in the prior section) from the “Ethical Decision-making” presentation in the “Assumptions I’ll Make About You” section of the first day’s class.

<https://ocw.smithw.org/mgt360/ethical-thinking-ws-short.pptx>

Additionally, make any reasonable counter-argument that your team feels is appropriate. You do this by using a different ethical *theory*. Again, draw upon the theories in our class “Ethical Decision-Making” reading (above).

## Conclusions

Provide any Recommendations you feel are appropriate. Also, no analysis is ever fully complete. Describe possible limitations of your team as appropriate. Might the team have suffered from one or more cognitive biases (a good list is available on Wikipedia). Could you have benefitted from a specialist from a different discipline? Essentially, what might be important but you know you are missing it?

Describe possible limitations of the data. What additional *quantitative* data might be useful for future work? Does it exist and you didn’t have time? Do you even know if the data exists at all?

---

## Sources for Non-Class Article

The sources for the articles for this class will be drawn from the following set:

- Newspapers (Dailies)
  - Los Angeles Times
  - New York Times
  - The (London) Times
  - Wall Street Journal

- Financial Times
- Magazines (Weeklies or Monthlies)
  - The Economist
  - Bloomberg BusinessWeek
  - Forbes
  - Fortune
  - Barron's

These sources embody a high degree of investigative journalism regarding business-related activities. To make a strong analogy, you want sources that emphasize authority, educational value, intent, originality, and quality. If you want to use something else, it must be approved by me in advance (it's possible, but have a back-up plan just in case).

### **Sources for LDC/UDC Concepts**

FIN 303 – Financial Management

<http://www.csun.edu/~jpd45767/303/fin303.html>

MGT 360 – Management and Organizational Behavior

<https://openstax.org/details/books/principles-management> (MGT)

<https://openstax.org/details/books/organizational-behavior> (OB)

<https://ocw.smithw.org/mgt360/textbook/carpenter.pdf> (MGT/OB)

MKT 304 – Marketing Management

<https://open.umn.edu/opentextbooks/textbooks/principles-of-marketing>

SOM 306 – Operations Management

<http://www.csun.edu/~vcmg0j3/SOM306.html>

In addition to your own textbooks and course materials, and the textbooks on reserve in the Oviatt Library Reserve Bookroom, the following textbooks may also be useful.

ACCT 220 – Introduction to Financial Accounting

<https://open.umn.edu/opentextbooks/textbooks/4>

ACCT 230 – Introduction to Managerial Accounting

<https://open.umn.edu/opentextbooks/textbooks/137>

ECON 160 – Principles of Microeconomics

<https://openstax.org/details/books/principles-microeconomics-2e>

ECON 161 – Principles of Macroeconomics

<https://openstax.org/details/books/principles-macroeconomics-2e>

BLAW 280 – Business Law I

[https://saylordotorg.github.io/text\\_foundations-of-business-law-and-the-legal-environment/](https://saylordotorg.github.io/text_foundations-of-business-law-and-the-legal-environment/)

Remember that each BUS 497a student has passed the BUS 302L “Gateway Lab” as well. While not textbooks, per se, the “top ten topics” and the concomitant review/study material for each of the six lower-division core subjects above make strong supporting and reference materials too.

<https://www.csun.edu/nazariancollegeadvisement/review-materials>

## **Possible Sources for Quantitative Reasoning**

A good start for data is the Business Databases available via the Library:

<https://libguides.csun.edu/bus/library-databases>

The FRED database is good for macro-level data. It’s run by the Federal Reserve Bank in St. Louis:

<https://libguides.csun.edu/az.php?a=f>

Data-Planet is well-organized (hierarchical) system of data that helps with various kinds of business, political, and social queries, with results that can be filtered by year and region.:

<https://libguides.csun.edu/az.php?a=d>

Remember too that for general, broad information, the U.S. Census (data.census.gov) can be quite helpful, especially for demographic data and trends (formerly known as the American Fact Finder):

<https://libguides.csun.edu/az.php?a=d>

Mergent Online also provides extensive information on companies and their competitors. It provides more firm information than what is required by the Securities and Exchange Commission (SEC) and can then therefore do industry-level analyses:

<https://libguides.csun.edu/az.php?a=m>

IBISWorld provides some of the best publicly available data on industries and markets. IBISWorld combines public- and private-data together for industry-by-industry analyses:

<https://library.calstate.edu/northridge/databases/alphabetical?alpha=I>

One CSUN-provided database for company stock prices is Y-Charts. Y-Charts is useful for accessing stock prices, among other things:



<https://library.calstate.edu/northridge/databases/alphabetical?alpha=Y>

Another CSUN-provided database is CalcBench. CalcBench is useful for accessing publicly-held financial statements. The source of this database is the SEC EDGAR system, however the web interface and usability is much improved over the standard EDGAR web interface:

<https://library.calstate.edu/northridge/databases/alphabetical?alpha=C>

The CSUN Library provides links to additional statistical data as well:

<https://library.calstate.edu/northridge/databases/subject/statistical-data>

Other sources of quality data that students have found useful in the past are:

<https://WSJMarkets.com/>

<https://google.com/finance>

<https://yahoo.com/finance>

<https://www.theharrispoll.com>

<https://www.statista.com/>

<https://www.axios.com>

## **Tips for Quantitative Reasoning**

Data Literacy, Analytics, and Analysis are a start but they, by themselves, are incomplete. Look for trends and patterns. What are you comparing and contrasting? Just solely in the area of organizational performance, there are at least five major kinds of performance evaluation:

### *Improvement*

Comparing current performance with past performance.

### *Comparative*

Comparing current performance to the performance of peers.

### *Goal*

Comparing current performance to the performance stated in a clear/explicit goal.

### *Ideal*

Comparing current performance to an ideal or perfect performance.

### *Stakeholder*

Comparing current performance to the expectations of one or more stakeholders.

You'll most likely rely on financial statements but there are other kinds of data too (e.g., marketing, operations, management). Recall the "Balanced Scorecard" (examples on Wikipedia). You must make a conscious decision to exclude one of these. Be prepared to justify your answer.

Just understanding data is difficult. *Data* is raw, *Information* is meaningful data, *Knowledge* combines explicit information with tacit information, and *Wisdom* is extraordinary insight or foresight. In your quantitative reasoning, can you distinguish between these in a clear way; that is, clear in the eyes of the decision-makers (audience)?

Which parts of your analysis are relatively *objective* (fact-based, education-based, primary-sources, first-person)? Which parts are relatively *subjective* (intuition-based, experience-based, secondary-sources, third-party)?

Are you *explaining* or *predicting*? Both are about understanding, however, explanations are of the recent past, and predictions are about the near future. Both are important but require different kinds of analytical techniques (the former, perhaps, hypothesis testing; the latter, perhaps, linear regression).

Are you *modeling* or *judging*? You develop models when you have some data; you use judgment when you don't have much data.

How have you maximized *rationality* and/or minimized *uncertainty* in the eyes of decision-makers? How have you leveraged *serendipity* and/or controlled *complexity*, again, in the eyes of the decision-makers?

Have you discussed *reward* and *risk* in the same context? One without the other will lead to weak results. Have you discussed *measurement* and *management* in the same context? One without the other is ineffective.