# Project Instructions

Team Formation:

Each team must comprise of 4- 5people

Dataset:

* You must use a real-world dataset and not fictious dataset
* Select a dataset from any of the following portal
  + <https://data.ontario.ca/en/dataset?keywords_en=Education+and+Training>
  + <https://data.worldbank.org/>
  + <https://datasetsearch.research.google.com/>
  + <https://www.who.int/data/gho/>
  + <https://registry.opendata.aws/>
* You can pick multiple dataset and use joins functions in SQL, Excel or Tableau to create a compelling business problem.
* Your dataset must contain at least 10 columns

Deliverables:

* + You should write a report, dashboard and a presentation

Presentation:

* Here’s an example of a sample report. “Making Brooklyn road safer” See attached e-conestoga
* You can get free slides from here: <https://24slides.com/templates/view/data-tables-graphs-charts/survey-results-powerpoint-template>

Project Analysis

How to write your report

The data analysis report has two very important features:

• It is organized in a way that makes it easy for different audiences to skim/fish through it to find the topics and the level of detail that are of interest to them.

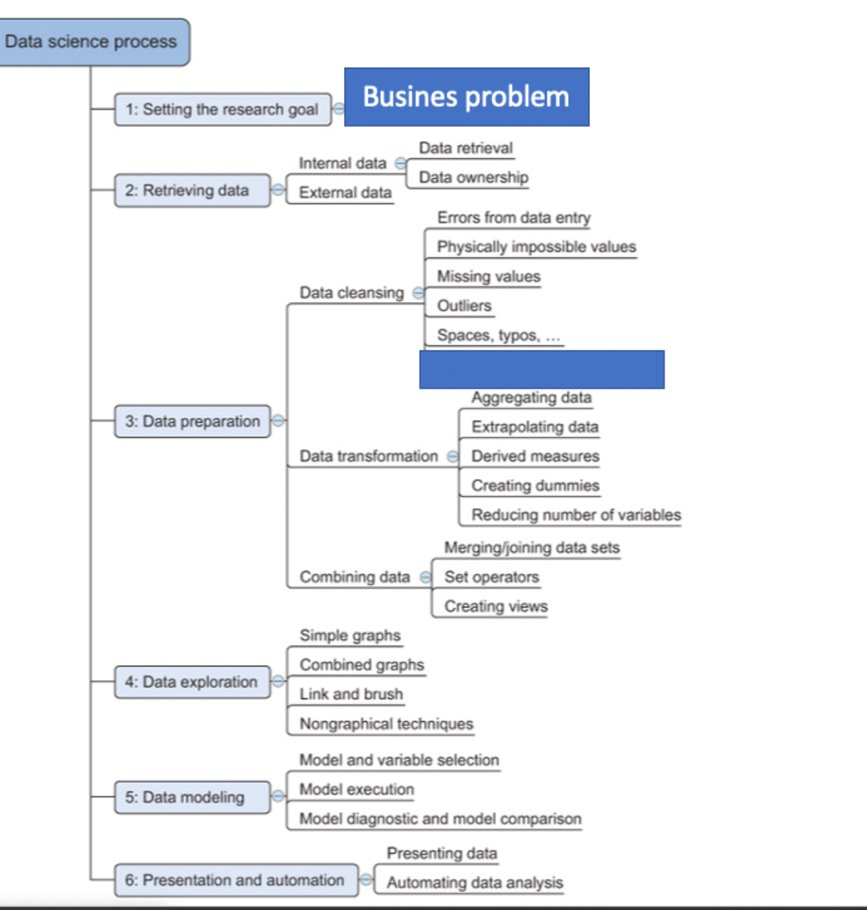
• The writing is as invisible/unremarkable as possible, so that the content of the analysis is what the reader remembers, not distracting quirks or tics in the writing.

Examples of distractions include:

* + Extra sentences, overly formal or flowery prose, or at the other extreme overly casual or overly brief prose.– Grammatical and spelling errors.
  + Placing the data analysis in too broad or too narrow a context for the questions of interest to your primary audience.
  + Focusing on process rather than reporting procedures and outcomes.
  + Getting bogged down in technical details, rather than presenting what is necessary to properly understand your conclusions on substantive questions of interest to the primary audience.

Structure of your report

Now let's consider the basic outline of the data analysis report in more detail:



**1. Introduction.** Good features for the Introduction include:

* + Summary of the study and data, as well as any relevant substantive context, background, or framing issues. For example:
  + The “big questions” answered by your data analyses, and summaries of your conclusions about these questions.
  + Brief outline of remainder of paper.
* For example: “Making Brooklyn safe”

2. Retrieving Data or Data source:

**4. Data Exploration: Have at least 5 insights in your report**

In this format there is a single Body section, usually called “Analysis”, and then there is a subsection for each question raised in the introduction, usually taken in the same order as in the introduction (general to specific, decreasing order of importance, etc.).

Within each subsection, statistical method, analyses, and conclusion would be described (for each question). For example:

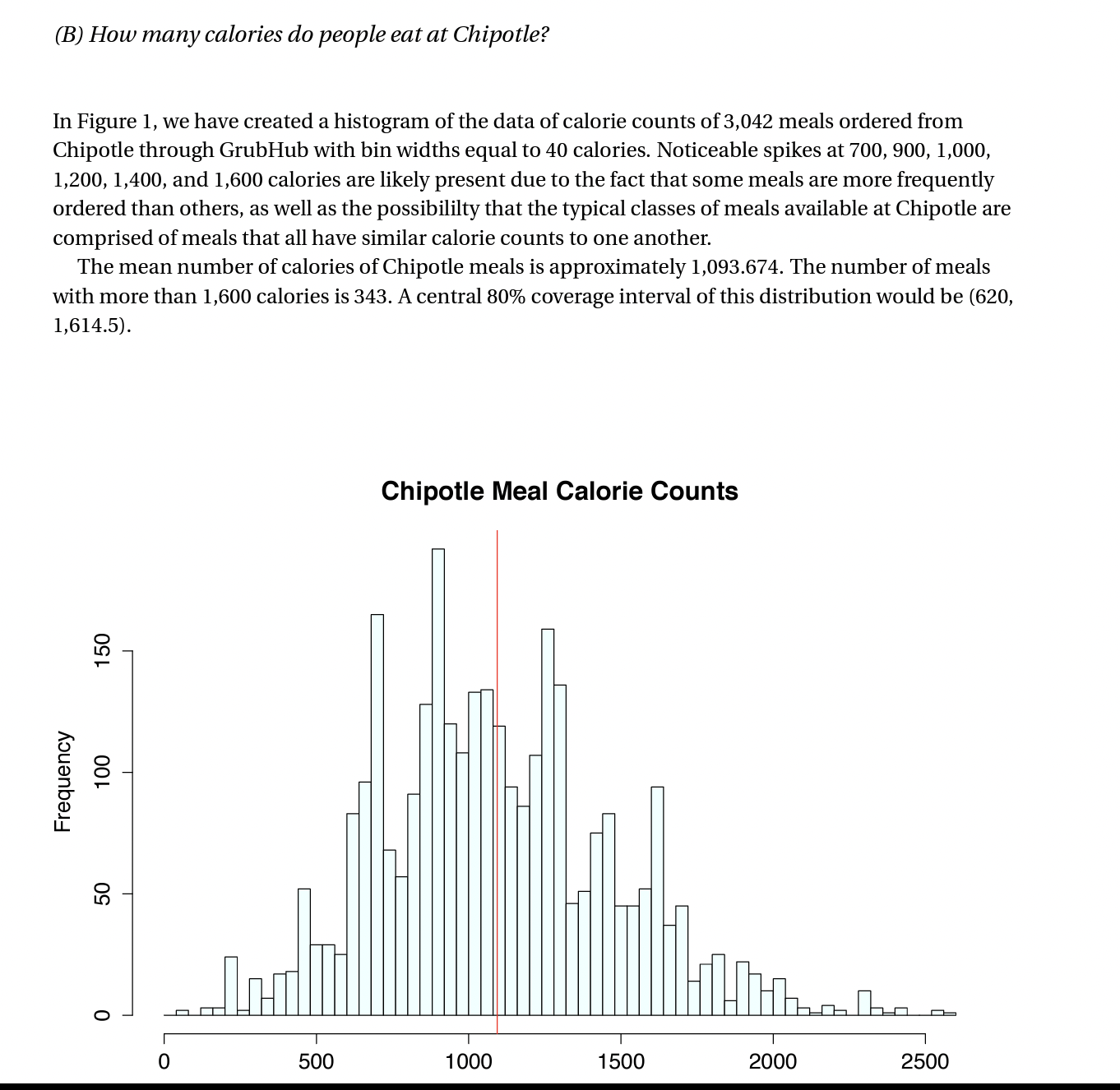
Here’s a sample of an analysis

2.1 Question or Metric 1: For example: How many calories do people eat at chipotle

Methods

Analysis

Conclusions

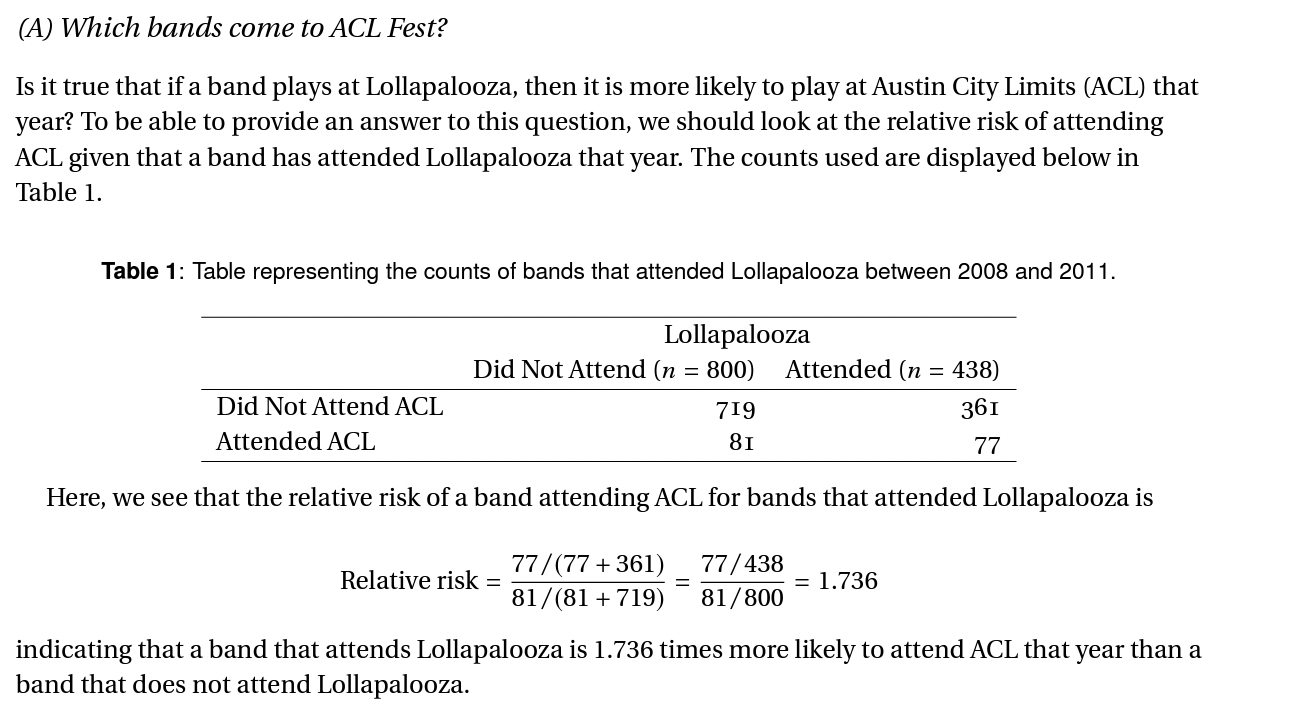


2.2 Question or Metric 2: For example: Which bands come to ACL Fest

Methods

Analysis

Conclusions



2.3 Question 3 or Metric 3: Exploring Ausin food critics

Methods

Analysis

Conclusions



Tip:

Other organizational formats are possible too. Whatever the format, it is useful to provide one or two well-chosen tables or graphs per question in the body of the report, for two reasons: First, graphical and tabular displays can convey your points more efficiently than words; and second, your “skimming” audiences will be more likely to have their eye caught by an interesting graph or table than by running text. However, too much graphical/tabular material will break up the flow of the text and become distracting; so extras should be moved to the Appendix.

**Appendix/Appendices.**

One or more appendices are the place to out details and ancillary materials. These might include such items as

• Technical descriptions of (unusual) statistical procedures

• Detailed tables or computer output

• Figures that were not central to the arguments presented in the body of the report

• Computer code used to obtain results.

In all cases, and especially in the case of computer code, it is a good idea to add some text sentences as comments or annotations, to make it easier for the uninitiated reader to follow what you are doing. It is often difficult to find the right balance between what to put in the appendix and what to put in the body of the paper. Generally you should put just enough in the body to make the point, and refer the reader to specific sections or page numbers in the appendix for additional graphs, tables and other details.