**PROJECT ONE**

The objective of this project is to use a variety of visual representations of data – including bar graphs, pie charts, line graphs, and stackplots – to identify and display key properties of data sets. This project will also require interpreting and drawing conclusions from such visual representations of data.

Additionally, a major component of this project is utilizing several important functions within standard spreadsheet software, such as Microsoft Excel. *Please note that developing proficiency with such software is required for the successful completion of this project and may take considerable time and effort if you do not already have extensive familiarity with such software.*

A spreadsheet file containing the data sets required for this project must be downloaded from the NAU MAT 114 course website. The file is located on the “Projects” tab of the course website and is referred to in this project as “the accompanying spreadsheet.”

All required charts/graphs **must** be created within a spreadsheet program. Handwritten or hand-drawn charts/graphs will not be accepted.

Answers to the following questions must be compiled into a single word-processing document (Microsoft Word, etc.), including necessary explanatory text and supporting charts/graphs. This final document must be neat and organized, with answers to individual problems clearly labeled. The method of submission of the final document (i.e. printed out, submitted electronically, etc.) will be determined by individual instructors.

Explanations of common visual representations of data are provided in Section 1B of the NAU MAT 114 course website. **Section 1B** also contains several videos demonstrating how to create such representations within the Microsoft Excel spreadsheet program.

This project must be submitted prior to the start of your Class Meeting during Week 3 of this semester.
ART FUNDING

1. The Texas Instruments Foundation recently donated more than $1 million to a variety of arts organizations in communities surrounding its north Texas headquarters. (View full report.)

Data provided in the accompanying spreadsheet indicates the amounts the Texas Instruments Foundation donated to the specific arts organizations.

a) Create a bar graph which displays the data set.

b) Create a pie chart which displays the data set.

c) Discuss which of the two visual representations you just created is a more effective representation of this data set.
FLAGSTAFF HOUSING

2. The website city-data.com gathers and displays a wealth of data regarding many metropolitan areas throughout the United States.

The graph/chart below, taken from city-data.com, details two key properties of home sales in the area code 86001 – _median home sale price_ and the _count of home sales_, both tracked quarterly over the last several years.

![Home Sales in Zip Code 86001](image)

a) Do you think this graph/chart provides an effective representation of this data set? Why or why not? Is there another common visual representation of data which you think would more effectively represent this data set?

b) While both quantities vary significantly over the time period under consideration, a general pattern seems to be present for the _count of home sales_ data – with one significant outlier. How would you describe the general pattern? In which quarter did the outlier occur?

c) In real estate, it is commonly assumed that when the price of homes increases, the number which will sell decreases, and vice versa. Does the data represented by the graph/chart above generally support or contradict this assumption? If possible, identify specific time periods to justify your answer.
3. The following page on city-data.com provides numerous visual representations of housing data for Flagstaff:


Scroll down the page until you come to a multi-bar graph entitled, “Rooms in houses/apartments in Flagstaff, Arizona.” Utilize the multi-bar graph to answer the following questions.

Note that the multi-bar graph allows you to switch between representations of data for *Owner-occupied* and *Renter-occupied* residences. Also note that the data displayed identifies the approximate number of houses/apartments in the given categories per 10,000 residences.

a) Is the number of Owner-occupied 5 room houses/apartments higher in Flagstaff or in all of Arizona?

b) Is the number of Owner-occupied 7 room houses/apartments higher in Flagstaff or in all of Arizona?

c) Does the number of rooms in a residence tend to be higher in Owner-occupied houses/apartments or in those which are Renter-occupied? Why do you think this is the case?

d) Does the number of rooms in an Owner-occupied residence tend to be higher in Flagstaff than in all of Arizona? Is the same true in Renter-occupied residences? Why do you think this is the case?
**ALTERNATIVE FUELS**

4. The number and variety of alternative fueling stations has grown significantly in recent years in the United States. Data provided in the accompanying spreadsheet was gathered by the United States Department of Energy and shows fueling station counts for the top 10 states.

   a) Create a column which calculates the *total number of fueling stations* in each state under consideration.

   b) Create a *pie chart* which displays the data for the *total number of fueling stations*.

   c) Create a *bar graph* which displays the data for the *total number of fueling stations*.

   d) Discuss which of the two visual representations you just created is a more effective representation of this data set.

   e) Create a row which calculates the *total number of stations* within each category.

   f) Create a *pie chart* which displays the data for the *total number of stations* within each category.

   g) Create a *bar graph* which displays the data for the *total number of stations* within each category.

   h) Discuss which of the two visual representations you just created is a more effective representation of this data set.

In addition to counts of fueling stations which are available within each state, the accompanying spreadsheet also provides the population of the states under consideration.

   i) Create a column which calculates the population of each state divided by the total number of fueling stations available within that state.

   *Note that this value indicates the number of people “per fueling station.”*

   j) Which three states are the “best” at providing alternative fueling stations?
STARBUCKS

5. Data in the accompanying spreadsheet tracks the number of Starbucks stores in the U.S., as well as the total count worldwide, between the years 1999 and 2012. Additionally, the spreadsheet provides the total annual revenue for Starbucks from 2001 to 2012.

A stackplot (also called an area graph) displays proportional relationships between quantities over time.

a) Create a stackplot which represents the relationship between Starbucks’ Total Number of Stores and Total U.S. Stores over the period from 1999 to 2012.

b) Based on the stackplot you just created, in which year does it appear that U.S. stores comprised the smallest percentage of total Starbucks stores worldwide?

c) Create a column which calculates the percentage of total Starbucks stores which are located in the U.S. in each of the years under consideration.

d) Create a line graph displaying the percentage data you just created.

e) Create a line graph displaying the Total Annual Revenue data.

f) Based on the line graphs you just created, discuss trends you see regarding the percentage of total Starbucks stores which are located within the U.S. and the company’s Total Annual Revenue.