**PROJECT TWO**

The objective of this project is to use a spreadsheet program to test correlation strength and to create and utilize mathematical functions which model relationships between pairs of quantities, especially relationships which demonstrate linear or exponential growth/decay.

Important definitions and helpful examples related to this project are provided in Chapter 3 of the NAU MAT 114 course website. Specifically:

- **3A** – independent/dependent variables, positive/negative correlation, correlation coefficient
- **3B** – comparing linear and exponential growth/decay
- **3E** – implementing functions using Microsoft Excel (video example)

A spreadsheet file containing the data sets required for this project must be downloaded from the NAU MAT 114 course website.

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. Explanations and discussions of answers must be in complete, grammatically correct sentences. The 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.

This project must be submitted **prior** to the start of your Class Meeting during **Week 9** of this semester.
**EAGLES**

A study was conducted on Admiralty Island in southeast Alaska using time-lapse photography to document Bald Eagle incubation, brooding (sitting on eggs to hatch them), prey deliveries and feeding. It provides a number of insights into male/female role differentiation including how roles and behaviors change based on environmental factors. While both male and female Bald Eagles brood, for example, the percentage of time the female broods during a given day appears to depend on several aspects of that day’s weather.

1. Data in the accompanying spreadsheet documents 19 days during the above study. For each day, the percentage of that day which was considered by scientists to be “sunny” was measured along with the percentage of that day during which the female Bald Eagle brooded (that is, sat on her eggs).

   a) Does it appear these quantities are positively related or negatively related? Why, do you think, these quantities would be related in this way?

   b) Create a scatterplot of the data.

      *(This is accomplished in Microsoft Excel by highlighting all of the data, clicking the “Insert” tab, and selecting “Scatter”).*

   c) Have your spreadsheet program create a linear regression trendline for the data, displaying the equation and the coefficient of determination ($R^2$) on the graph.

      *(This is accomplished in Microsoft Excel by right-clicking on a data point in your scatterplot, selecting “Add Trendline…”, and selecting “Display Equation on chart” and “Display R-squared value on chart”).*

      **Include a copy of this scatterplot in your report.**

   d) Find the correlation coefficient $R$ for your linear regression trendline. Does that value indicate a strong correlation or a weak correlation?

      *Note: You must calculate and discuss $R$, but the spreadsheet program displays $R^2$.*

   e) Based on the linear regression trendline you just generated, what percentage of a day would you expect the female Bald Eagle to brood if 50% of that day was considered “sunny”? *Please show calculations to support your answer.*

   f) Based on the linear regression trendline you just generated, what percentage of a day would you expect the female Bald Eagle to brood if 100% of that day was considered “sunny”? *Please show calculations to support your answer.*

   g) Have your spreadsheet program create an exponential regression trendline for the data, displaying the equation and the coefficient of determination ($R^2$) on the graph. Find the exponential correlation coefficient ($R$).

      **Include a copy of this scatterplot in your report.**

   h) Based on the two correlation coefficients you have just calculated, which type of function (linear or exponential) appears to be a better fit for this data?
The owner of a local skateboard and snowboard shop has kept careful records of the money she has spent on advertising – both in a variety of print formats (newspapers, magazines, etc.) as well as online – since the shop opened in January, 2008. In addition, the owner has tracked sales figures in each of the following categories:

- **Clothing (General)** – t-shirts, button-up shirts, pants, casual outerwear, socks, etc.
- **Clothing (Technical)** – purpose-specific clothing designed for snowboarding (high-performance pants and jackets, snowboard boots, etc.)
- **Shoes**
- **Accessories** – hats, wallets, watches, sunglasses, etc.
- **Skateboard equipment**
- **Snowboard equipment**

Data in the accompanying spreadsheet accounts for advertising and sales figures as reported in two-month intervals from 2008 to 2014.

2. Analyzing the strength of the correlation between dollars spent in advertising and sales figures in various categories will help to gauge the effectiveness of the owner’s advertising efforts.

   a) Create a scatterplot for each pair of quantities below. Have your spreadsheet program create a linear regression trendline for each scatterplot, displaying the coefficient of determination ($R^2$) on the graph.

   - Clothing (General) sales vs. Print Advertising
   - Accessories sales vs. Print Advertising
   - Skateboard Equipment sales vs. Print Advertising

   **Include each of these scatterplots in your report.**

   b) Find the correlation coefficient $R$ for each linear regression trendline you just generated.

   c) Sales in which category of merchandise analyzed above have the strongest correlation to the amount of money spent on print advertising? Would you categorize this correlation as very strong or only moderately strong?

   d) Sales in which category of merchandise analyzed above have the weakest correlation to the amount of money spent on print advertising?

   e) Create a scatterplot for each pair of quantities below. Have your spreadsheet program create a linear regression trendline for each scatterplot, displaying the coefficient of determination ($R^2$) on the graph.

   - Clothing (General) sales vs. Online Advertising
   - Accessories sales vs. Online Advertising
   - Skateboard Equipment sales vs. Online Advertising

   **Include each of these scatterplots in your report.**

   f) Find the correlation coefficient $R$ for each linear regression trendline you just generated.

   g) Sales in which category of merchandise analyzed above have the strongest correlation to the amount of money spent on online advertising? Would you categorize this correlation as very strong or only moderately strong?
h) Sales in which category of merchandise analyzed above have the *weakest* correlation to the amount of money spent on online advertising?

In addition to analyzing sales in the given categories, it will be helpful to test the strength of the relationship between *total sales* and the investments in advertising.

3. a) Create a column adjacent to the given table which calculates the *Total Sales* in all of the categories for each time period.

b) Create a scatterplot for each pair of quantities below. Have your spreadsheet program create a *linear regression* trendline for each scatterplot, displaying the coefficient of determination (*R²*) on the graph.

   *Total Sales vs. Print Advertising*

   *Total Sales vs. Online Advertising*

   **Include each of these scatterplots in your report.**

   c) Find the correlation coefficient *R* for each linear regression trendline you just generated.

   d) Which form of advertising – print or online – has the stronger correlation total sales.

   e) Based on all of the analysis you have completed so far, what recommendation would you make to the shop owner regarding future spending on advertising?
**Waste Management**

Growing cities face the challenge of efficiently disposing of or recycling the solid waste generated by their increasing populations. Fortunately, thanks to advances in recycling technologies, the amount of waste which can be reused or recycled increases every year.

The following questions will model the growth in the population of the town of Marana, Arizona and estimate the demand for waste management services generated by this population.

4. a) The town of Marana experienced dramatic growth beginning in the 1990s. If the population of Marana was 3,316 in 1990 and growing at a rate of 14.99% per year, create a table of values indicating the population’s growth through 2006.

b) After the period of rapid growth described above, the growth of Marana slowed significantly. Town officials estimate growth of only 2.869% for the next several years. Assuming this lower growth rate began in 2006-2007, continue your table of values to project the population of Marana through 2014.

c) It is estimated that each individual generates approximately 1424 pounds of solid waste each year. Create a new column of values which indicates how many pounds of solid waste have been generated by the population of Marana from 1990 to 2014.

Advances in recycling technologies have reduced the amount of solid waste which goes to a landfill each year. In a typical municipality, the amount of solid waste which could be recycled in the early 1990s was approximately one million pounds per year. Depending on the municipality, this amount may have grown by as much as 600,000 pounds per year since.

5. a) In 1990, the town of Marana had the ability to recycle 1,110,000 pounds of solid waste per year. This capacity increased by a fixed amount of 509,850 pounds per year until 2002, when Marana introduced several recycling initiatives. Since 2002, Marana town officials have boasted an increase in recycling capacity of approximately 6.5% per year. Create a column of values indicating how many pounds of solid waste have been recycled by Marana from 1990 to 2014.

b) Assuming the number of pounds of solid waste which end up in a landfill is the total solid waste generated minus the number of pounds recycled, generate a column of values indicating how many pounds of solid waste end up in a landfill due to the population of Marana from 1990 to 2014.

c) Create a multi-line graph displaying the total solid waste generated and the solid waste which ends up in a landfill due to the population of Marana from 1990 to 2014.

**Include your completed table of values and your multi-line graph in your report.**