Project Name: Shoes-GALORE

Team Members:

* Danelle Cook
* Jason DeGrace
* Jennifer Long
* Mark Alford
* Rashaye Harris
* Avondre Henderson
* Theadora James

Project Description/ Outline: For the basis of this analysis our group will be looking at two datasets that look at Men’s and Women’s shoe sales. We’ll be taking a deep dive into comparing sales by brand. Distribution, and gender – looking at max, min, and average prices.

Looking at the Women’s Shoes Data Set

Research Questions to Answer:

* Introduction
* (Together)What is the impact of women on shoes and what kinds of shoes are being produced? What category of shoe is most widely made for women? Boot, sandal, heels, etc. (Pie chart)
* (Avondre, Jennifer) What are the top colors and the average price points of those colors? (Pie chart)
* (Rashaye, Mark) Out of the shoes being produced, what are the top brands that are produced? Based on the highest amount of individual skus and highest price point. (Bar chart)
* (Thea)What is the average price point of the top brands? (Box plot)
* (Jason, Danelle) Does shoe type affect pricing? (Line graph or scatter)
* Analysis

Datasets to be used:

<https://data.world/datafiniti/womens-shoe-prices>

Rough Breakdown of Tasks:

* Pull both of the CSV files into a jupyter notebook.
* Clean the data – be sure to remove any blank data cells (and/or replace them with 0)
  + When cleaning the data be sure to make sure that everything is of the same item type.
* Create a dataset for the newly cleaned data.
* Create Data visualization
* Report on findings and present to the class.

Notes for the Presentation

* We can explain the limitations of the data at the end. We can’t connect to an API in this case because we don’t have any geographical data.