**Slide 3 Script:**

* Hi
* Queen Bees
* Data Science Course
  + Entity
  + Woz-U
  + Learning Source

**Slide 5 Script:**

* Sara Slocum
* Santa Monica and Miami Dade Colleges
  + Finance, Music, & Performing Arts
* Lots Real work experience & Professional certifications
  + S6, 65 & Life & Health
  + R.E.
  + Cycling Instruction & Group Fitness
  + Lifeguarding, CPR, & First-Aid
* Currently self-employed
* Previously Investment advisor & retail store manager.
* Creative, empathetic leader
* Healing Manuka Honey & Signature Honey Latte

**Slide 9 Script:**

* Two Datasets
  + Kaggle
  + USDA
    - Govt. Information
    - Supply, Production, Trade, & Consumption
    - Worldwide
    - Yearly by Country 1960s-2022
    - No Honey info
  + US Honey Production
    - By State, per year 1998-2012
    - Added state totals to get yearly totals
    - Appended to USDA df
  + Bulk of Analysis
    - US Production values 1998-2012 as they relate to honey

**Slide 13 Script:**

* Left Table: Avg. Yearly Production Top Producing Crops
  + Yearly Ave Metric Tons
  + % of total US production yearly aves
  + Top 3 Outproduce all others
    - Corn 54%
    - Milk 16%
    - Wheat 11%
  + Next crop, Chicken falls to 2.7% of total
  + Other commodities at bottom, remaining 10%
  + What % of total is Honey?
    - .01%
    - Still not small – 78,920 MT
      * Weight equivalent of over 13,000 elephants
* Pie Chart: Total Sum of all Production 1998-2012
  + Similar ratios to the yearly ave
  + Data showed dom. Consumption paralleled production. With some exceptions
    - Corn, Dairy, Wheat top consumed
* Modern Pyramid chart
  + US may want to adjust production & consumption to match
  + More nutritious vegetables – corn not nutritious

**Slide 17 Script:**

With this background information in mind we’ll move into our analyses

**Slide 18 Script:**

* 6-step process
  + 1) Corr Matrix
  + 2) Singled out commodity production correlating highly to honey production
    - 65% or higher
  + 3) Linear Regression on high correlators
    - Change over time 1998-2012
    - Relationship w. honey during time period
  + 4) Linear Equations
    - Define ea. Commodity’s relationship w/ honey
    - To use for predictive forecasting
  + Post Analyses: Can we make predictions about honey production for future?
    - Honey to 2012, other commodities to 2022
  + 5) ARIMA
    - 2023 Forecast values on high correlators
    - Plug in to linear equation & solve for honey
      * One prediction for each high correlator
    - Create df & scatterplot of honey prediction values
  + 6) K-Means cluster analysis
    - Compare plots to eachother
    - Observe the spread of prediction points
    - Make conclusions about 2023 forecasts

**Slide 23 Script:**

* Nuts & Oils Group:
  + Almonds, Filberts, Pistachios, Walnuts, Peanut Oil, Olive Oil, Rapeseed Oil, Cottonseed Oil, & Soybean Oil
  + Quick question, what are filberts?
    - Filberts = Hazelnuts, one of my favorite nuts
* Correlation matrix, four high correlators:
  + Almonds at 81%,
  + Walnuts at 74%,
  + Rapeseed Oil at 73%, and
  + Soybean Oil at 67% correlation
* Linear trend plot
  + Honey Production Values on x axis
  + Commodity production value on y
    - Individual axis for soybean oil
      * To make plottable
      * Production values exceed the other commodities by tens of millions
  + High correlators all negatively correlated with honey production
    - As Honey production increases, the production values of these four commodities decrease.
    - Interesting find: fluctuation in production values between the commodities all occur similarly at the same levels of honey production.
      * Refer to visual
      * Something I’d like to look further into given more time.
* Linear Equations of regressions to honey
* To wrap up the analyses on Nuts & Oils,
* During this time period, Honey accounted for variance in:
  + Almond production by 63%
  + Walnut production by 52%
  + Rapeseed Oil production by 50%
  + Soybean Oil production by And 41%
* Sonya will take it away for our miscellaneous commodity analysis

**Slide 30 Script:**

Last but not least, we want to give some disclaimers about the analyses we have done, and touch on where we would go next to further this study.

* Limitations
  + The three biggest limitations: Time, Data, and Knowledge
  + Extensive dataset
    - So many facets to look at not enough time
    - Deep into wrangling, we found ourselves questioning accuracy of some entries.
      * Certain things didn’t add up or make sense.
    - We also realized some of the questions we wanted to answer required analyses beyond what we learned in the course.
    - Assumption issues
      * Although small could affect the results
    - Other factors we didn’t have data for:
      * Weather, Regulations, etc.
* Next Steps
  + Finding Better Data to improve data set
    - Although directly from the USDA.gov website, want to find other agricultural data elsewhere
      * More comprehensive and thorough
      * To compare to or update our existing data.
      * Perhaps from some other official sources on global agricultural production and trade
  + Fixing Assumption Issues
    - Homogeneity issues common across all regressions
    - Removing outliers
      * Massive wrangling task considering size of dataset
  + Reading the many 400 page texts Professor Raetano gave us on ARIMAs
    - To develop deep understanding of them
    - Be able to run them all manually,
      * Fine-tuning each parameter
      * Improve the R2 scores we got (accuracy)
  + Look into other factors effecting production values
    - Incorporate them for precision in forecasts
    - To explain causation of fluctuations & movement
  + Exploring massive amounts of global data on Trade & Consumption in USDA dataset

Although we still have plenty to do with this existing project, we already have an idea in mind for our next project. Please stay tuned for the Queen Bee’s next project (drum roll please)…

**Slide 31 Script:**

How could the information we have based on our data and analyses impact futures pricing?

That wraps up our presentation of US Honey Production, as it relates to other commodities. We would like to dedicate the last few minutes to Q & A to answer any questions you may have.