**Data Analysis Project #3**

**Introduction**

In this project you will continue to combine your previous training in Python and SQL along with new Tableau tools to create a final product for your portfolio. The Tableau objectives below are part review and part introduction to new features and concepts.

**Tableau Learning Objectives with Review(Modules 8-10)**

|  |  |
| --- | --- |
| Calculated fields with built-in functions | Table Calculations |
| Sets, Parameters, Group | Trend Analysis |
| Reference Bands and Distribution | Box Plot, Bins |

**<**

**Definitions**

**Geographic Analysis -**  An analysis of trends within or across groups defined by their geographic region.

**Temporal Analysis –** An analysis of data based on specific period of time or change over time.

**Intuitive Analysis -**  An analysis of data across groups based on logical explanations and/or behavioral patterns

**Dask-**  [Dask Python](https://docs.dask.org/en/stable/dataframe.html) library that works with large datasets, converting them to a large parallel DataFrame consisting of many smaller Pandas Data Frames. The data is split along the index.

* To install the library: pip install dask
* To import: import dask.dataframe as dd
* Sample import:

filepath = *‘/Path/to/file.csv’*

dsk\_df = *dd.read\_csv(filepath, low\_memory = False)*

dsk\_df = *dsk\_df.compute()* #converts the data frame to a Pandas data frame

**Project Objective**

In this project, you will work with the Chicago Police Department (or another city police department) to identify trends in crimes across the city. Think of ways to view the data, e.g. data occurring over time or by geographical locations.

**Project Tasks**

**Methodology**

**Ask Questions to Build Your Data Model.**

1. For Chicago data, use this website: [Chicago Crime Data 2001-Present](https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2)
   1. Additional mapping files to wards/neighborhoods are posted in Github a repository
   2. Additional links will be posted in Slack Resource channel
2. Make a list of 7-10 questions to ask about the data.
3. Use descriptive statistics to help explain your findings. Consider the distribution of the data. Think about possible trends
4. **Submit the list of questions in place of Module 8 challenge.**

**Github**

1. Create a new repository called ‘<City> Crime Data Analysis’
2. Save all of your files in your repository
3. **Submit the link to your repository in Module 9 challenge.**

**SQL and Python**

1. You can import the file directly into Python or a SQL database to review the data.
2. Connect to the database from Python to import your model.
3. Use the questions to build your analysis of the data in Python,
4. Double check your numbers in SQL.
5. Practice creating a function in Python
6. Build 2-3 charts in Python.
7. Export data to Excel or a CSV file
8. **Submit your Python code with charts and CSV file or your Github Repository link in Module 9**

**Tableau**

1. Create a new Tableau workbook, named ‘<City> Crime Data Analysis’.
2. Connect your data file then save your workbook. This will force you to log in. Make sure to save your work periodically or you will lose everything!
3. Confirm that all columns appear in the Data Source
4. **Data Cleaning**
   * 1. Think about where you can use a calculated field
     2. Are there any values that need to be replaced? How can you do this?
     3. Think of where you can apply a filter
5. **Preliminary Dashboard Charts** (You can change this. It’s designed to get you started.)
   * 1. Create the following worksheets:
        1. A bar chart
           1. Trend line
           2. Table calculations (Your choice)
        2. A box plot
        3. A map
        4. A scatterplot
           1. Include a trend line
        5. 2 Worksheets:
           1. Total Arrests. Name the worksheet ‘Total Arrests’.
           2. Total
6. **Dashboard Design**
   * 1. Use PowerPoint or some other design tool to plan your dashboard layouts before building. Explore the Tableau Public Gallery for ideas. You can also change the measures and charts if you want. Think about the color scheme you want to use. Use this site to help, <https://www.toptal.com/designers/colourcode>.
     2. Create a Dashboard
        + 1. Adjust the width to 1200
          2. Delete the container
          3. Drag text object to the dashboard

Change the object to ‘floating’

Change font type

Increase the font size to 48

Change the background color

Change text color

Make the bold

* + - * 1. Use your mock-up to build your dashboard. Make sure you:

Include filters that updates all of the data in the dashboard when an item is selected

Publish your final dashboard to your Tableau Public site

* + 1. **Submit the link to your dashboard in Module 10.**