CTEC 298-101 Spring 2021 1st 8w March 4, 2021

1. **Define the terms below using the definitions found in the tutorials that were provided to the class. Each answer will consist of the definition and the tutorial where it was found.**

Jupyter Notebook – **an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text.**

Cell – **an area on a spreadsheet where data can be entered**

Kernel – **the essential center of a computer operating system**

Ipynb file – **is a notebook document created by Jupyter notebook**

JSON **– is a standard text-based format for representing structured data based on JavaScript object syntax**

Metadata – **a set of data that describes and gives information about their data**

Code cell - **allows you to edit and write new code, with full syntax highlighting and tab completion**.

Markdown cell – **displays text which can be formatted using markdown language**

Matplotlib Plots – **is a plotting library for the Python programming language and its numerical extension Numpy.**

1. Describe each of the 6 plots that we have used in Matplotlib

**Line Graph, Bar Graph, Histogram, Scatter Plot, Stack Plot, Pie Plot**

1. What is a panda in this class?

**Panda is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.**

1. How are Pandas related to Excel?

**It can be used as function arguments and return types for Excel worksheet using the decorator xl\_func**

1. Pandas are used for what is the major data science concept.?What is a dataframe

**Panda is used for data manipulation and analysis. DataFrame is a 2-dimensional labeled data structure with columns of potentially different types.**

1. Compare Matplotlib and Tableau for creating plots.

**Tableau helps people see and understand data. Tableau can help anyone see and understand their data. Connect to almost any database, drag and drop to create visualizations, and share with a click. Phython A plotting library for the Python programming language. It is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. It can be used in Python scripts, the Python and IPython shells, the Jupyter notebook, web application servers, and four graphical user interface toolkits.**

1. What types of input files can be used for each

1. What is a dataframe? Give an example.

**Dataframe is the most common Structured API**

1. Name 3 statistics concepts that can be used with dataframes?

10 Describe 3 techniques for creating a dataframe using Pandas>