CSSE 386 – Data Mining with Programming
Rose-Hulman Institute of Technology

Quiz 01 Solutions

Name (Print):_____   Section:_____

1. Who is a pioneer of Exploratory Data Analysis?.

    a.) Marc Zuckerberg

    b.) **John Tukey**

    c.) William Cleveland

2. When Python and R were first released?

    a.) 1980

    b.) 1960

    c.) **1990**

    d.) 2000

3. What are the three V of Big Data? Select three!

    a.) Value

    b.) **Volume**

    c.) Variable

    d.) **Velocity**

    c.) **Variety**

4. What are the steps of the Data Mining Cycle?

    1.) **Gather**

    2.) **Clean**

    3.) **explore**

    4.) **Model**

    5.) **Interpret**

5. Which Processing Unit will you use for working with AI and cloud computing?

    a.) CPU

b.) GPU

c.) **T4 GPU**

d.) TPU

6. Which library will you use for visual analysis?

   a.) **matplotlib**

   b.) pandas

   c.) os

   d.) numpy

7. Describe the difference between Data Mining and Machine learning in your own words:

   DM - discover insights, ML - design algorithms

8. What is the alternative name for columns in the data frame?

   a.) instances

   b.) **features**

   c.) dataset

   d.) matrix

9. Decision tree is the method developed from which field?

   a.) Statistics

   b.) **ML**

   c.) AI

10. Linear and Logistics regressions are the methods developed from which field?

    a.) **Statistics**

    b.) ML

    c.) AI