Multiple choice concept questions:

Q1. What is known as the grandfather of all recursive date structures?

1. Prefix tree
2. Binary tree
3. Binary prefix
4. Suffix tree

Answer: (b)

Q2. What is the key factor in connecting to audience for a data scientist?

1. Power
2. Asking questions
3. Details
4. Story telling

Answer: (d)

Q3. Which one is an example of basic models of time series analysis?

1. Linear models
2. Arima model
3. Auto regression
4. Black sholes

Answer: (a)

Q4. What kind of time series analysis is the Recurrent Neural nets?

1. Advanced models
2. Multivariate
3. Basic models
4. ML models

Answer: (b)

Q5. Which one is not a function to draw different shapes?

1. Circles
2. Ellipses
3. Polygons
4. Skimage

Answer: (d)

Q6. What is the main alternative to the vector data types?

1. Rasters
2. Polygons
3. Points
4. Lines

Answer: (a)

Q7. Which sentence is true?

1. In Gall\_peters projection, sizes are not equal to the reality, but shapes are the same.
2. In Mercator projection, sizes are equal to the reality, but shapes are distorted.
3. In Gall\_peters projection, sizes are equal to the reality, but shapes are distorted.
4. In Mercator projection, there is a pleasant balance between shape and size accuracy.

Answer: (c)

Q8. What is the Backpropagation?

1. It is a technique to train a neural net
2. It is an algorithm that learns to identify patterns in data
3. Analyze the implications of results
4. Send signal to the neural networks

Answer: (a)

True-False Choice

Q1. Asymptotic analysis characterizes time and space complexity of method.

1. True
2. False

Answer: T

Q2. Fundamental issue facing all programmers is implementing their own data types.

1. True
2. False

Answer: T

Q3. A common solution for contiguous memory problem in early computer languages was queues.

a. True

b. False

Answer: F (linked lists)

Q4. The bridge between the data and humans is data visualization.

1. True
2. B. False

Answer: T

Q5. Bar plots and box plots are multivariate non\_graphical methods.

1. True
2. B. False

Answer: F (graphical)

Q6. Vet and test your ideas is a goal for story boarding and sketching.

1. True
2. B. False

Answer: T

Q7. Polygons are dimensionless, just a location in space.

1. True
2. False

Answer : F

**Example Short Answer Question**

Q1.

We are going to choose an image and do some changes on its color.

#Import necessary modules:

From skimage import io, color

#open an image

img = io.imread(“image.png”)

#convert RGB to Gray:

gray = color.rgb2gray(imag)

#show the image

io.imshow(gray)

io.show()

Q2.

Download data from Kaggle, and read it into python and save it as panda data frame.

Create histogram for all the numerical variables. Explain your results. (Do they have normal distribution?)

Q3.

Using the data set in the previous step, create a scatter plot between a pair of numerical variables. Do you see any pattern among them?

Q4.

Using the data set from Q2, choose any two variable of interest, and build a linear regression model. Estimate your parameter.

Q5.

Repeat Q3, but first split your data into 70% training and 30% test. Evaluate your model performance using test data.

Data Set Mini

Using the data set in Q1. Normalize all the numerical variable.

Then calculate the mean and standard deviation to verify that the data is normalize.