DATA-101 IN-CLASS ASSIGNMENT OCTOBER 27, 2022

1. (3 points) Suppose you have been given a fair coin (P(head) = P(tails) = 0.5) and you want to find out the odds of getting heads. Which of the following option is true for such a case?
   (a) Odds will be 0
   (b) Odds will be 0.5
   (c) Odds will be 1
   (d) All of the above
   (e) None of the above
2. (3 points) Is logistic regression a supervised learning method?
   (a) True
   (b) False
   (c) It depends
   (d) None of the above
3. (3 points) Is Logistic regression mainly used for Regression?
   (a) True
   (b) False
   (c) It depends
   (d) None of the above

Consider the `Customer_Churn.csv` datafile. Each row represents a customer, each column contains customer's attributes described on the column Metadata. The data set includes information about:

- Customers who left within the last month, the column is called Churn.
- Services that each customer has signed up for phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.
- Customer account information: how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges.
- Demographic info about customers: gender, age range, and if they have partners and dependent

4. **In Python**, answer the following:
   (a) (3 points) Using the `pandas` library, read the csv file and create a data-frame called `churn`.
   (b) (3 points) Using the `numpy` library, create a variable called `Churn_numb` that takes the value of 1 when `Churn = Yes` and 0 when `Churn = No`.
   (c) (5 points) Using the `statsmodels.formula.api` library, build a logistic regression model, in which `Churn_numb` is the target variable, and `tenure` and `MonthlyCharges` are the input variables. Predict the likelihood of churn of a customer with `tenure` equal to 20, and `MonthlyCharges` equal to 80.