After you run a campaign and get back purchase decisions, you'd like to figure the probability of a customer from your database ordering based upon four continuous variables. Which of the following techniques will you use?

|  |
| --- |
|  |

|  |  |
| --- | --- |
| **A. Correlation** |  |
|  |  |
|  |  |
|  |  |

You might have been asked to choose the closest to correct model for the response variable: (Hint: you'll need to recode Response from Y/N to something else) Using the p-values of the coefficients in the result we can reduce the linear equation to a simpler form:

|  |
| --- |
|  |

* **E. Response = 5 -.0644 \* Age**

In BigClass, The correlation between AGE and WEIGHT is?

|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | |  |  | |

* **B. No correlation**

Propose a multiple-choice question covering the materials from the course and provide at least four answers to it. Very good questions will earn 2 points bonus.

Why is NYU SPS not a good representative sample of New York City?

|  |
| --- |
|  |

**B. New York is diverse, made up of many types of people of all ages and college graduation status**

We need to reach people living in a specific zip code in Brooklyn, primarily male, Hispanic, 35 to 54. What would a representative sample of this population look like?

|  |
| --- |
|  |

* **A. Older than our target audience**

Analyzing the attached chart what kind of correlation do you see?

|  |
| --- |
|  |

* **A. Positive correlation**

You have run a marketing promotion that, in the past, has returned 14% response rate from the targeted group. You've recently decided to tweak the campaign to see if you get a higher response rate, so you run a campaign using both versions of the promotion, and evaluate it using chi-squared analysis. Of 4790 people promoted under the new model, 730 responded, a  Blank 1. Fill in the blank, read surrounding text. % response rate when rounded to the nearest whole percentage; meanwhile 750 of 5210 people promoted with the old promotion responded, giving you a  Blank 2. Fill in the blank, read surrounding text. % response rate, when rounded to a whole percentage.  
  
You decide that this  Blank 3. Fill in the blank, read surrounding text. (write is or is not) a statistically valid result.

The primary goal in establishing a marketing database is:

|  |
| --- |
|  |

* **C. Helping the organization achieve an objective such as increasing profits by tracking campaigns**

CMO Grace Liu wants to aggregate data together because individual data points are too few by themselves to generate meaningful information. What is the CMO creating?

|  |
| --- |
|  |

* **D. Longitudinal variables**

Using the attached chart, the tab for question 23, if Chou En Lai figured out that the goal of the promotion is to generate a 6% profit after overhead and the resulting response rate had to be 3.7%, which names can Chou promote profitably?

|  |
| --- |
|  |

* **A. None**

Choose the closest to correct model for the response variable: (Hint: you'll need to recode Response from Y/N to something else)

* **D. Response = 5 -.0644 \* Height - .00007 \* Age**

**Q13-34**

From the attached Excel spreadsheet go to the Excel tab for Q25 and Q26. Determine which person had the highest order to promotion ratio:

|  |
| --- |
|  |

* **D. Aggata**

Use the BigClass.xls file associated with this exam. Open it in Excel and compute the correlations asked for. What is the correlation between the AGE and HEIGHT variables?

|  |
| --- |
|  |

* **A. .60826**

Which of the following is not a valid reason for purging or removing customers from your marketing database?

|  |
| --- |
|  |

* **C. Customer only ordered during the past 6 month promotional period**
* 

From the Excel spreadsheet. tab for question 28, 29, 30: what is the cumulative response rate for those people with a response rate of 7.5% or greater?

|  |
| --- |
|  |

* **A. 29%**

In BigClass, Why is it not possible to compute the correlation between AGE and GENDER?

|  |
| --- |
|  |

* **B. GENDER is a categorical variable**
* Using the attached chart, go to the tab for question 22: assuming the breakeven for hot fire chilies in a jar is a response rate of 2.50%, which names could a product manager profitably promote?

|  |
| --- |
|  |

* **B. No age available and 31 to 40**

Working with categorical data requires some special handling techniques, of which we've used a few. Which of the following techniques are valid when used correctly with categorical data?

|  |
| --- |
|  |

|  |  |
| --- | --- |
|  |  |
|  |  |
| **C. Chi-Squared** |  |
|  |  |

A key assumption of CRM is:

|  |
| --- |
|  |

* **B. A focus on immediate sales volume leads to better CRM**

What is the best way to do an Nth select on Excel so you get about N rows but every row has an equal chance of being selected?

|  |
| --- |
|  |

* **A. Assign a random number between 1 and N inclusive to each row, and choose every row that gets a 1**

In BigClass, the correlation between AGE and WEIGHT makes sense since

|  |
| --- |
|  |

* **A. There is no discernible correlation between these two variables**

Suppose, instead of 95% confident as in Question (Go to the sample size calculator: https://www.surveysystem.com/sscalc.htm You want to do a representative sample for a political poll on the population of China, given as 1400000000, with 95% certainty that your poll result will be within 2% correct [your confidence interval]. How large will your sample size be?) you want to be 99% confident. How large does your sample size have to be?

|  |
| --- |
|  |

* **C. 2401**

Which of the following is a standard database maintenance routine?

|  |
| --- |
|  |

* **A. House-holding the customer file**

CEO Launa Lee is concerned about a biased result in her analysis, thus she needs to do what?

|  |
| --- |
|  |

* **D. Pull names that are independent of the final result**

Based on the Ratio Variable in the attached Excel spreadsheet under tab Q25, 26, who would you be least likely to promote in the future based on just orders?

* **D. Beverly**

In BigClass, The correlation between the AGE and WEIGHT variables can be considered a strong correlation.

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| |  | | --- | |  | | False | |

Using the attached spreadsheet, go to tab for question 26 for this question. For reasons you can’t believe, CMO Martin wants you to promote the two people who have the highest combined orders. Which two people would that be?

* **C. Aggata and Chu**

Go to the sample size calculator: https://www.surveysystem.com/sscalc.htm You want to do a representative sample for a political poll on the population of China, given as 1400000000, with 95% certainty that your poll result will be within 2% correct (your confidence interval). How large will your sample size be?

|  |
| --- |
|  |

* **D. 2401**

From the attached chart in the tab entitled Q28, 29, 30, compare cells C1, C6, C7, C13, C14, C19, C20. Determine which cell has the highest index number. What does it mean? Select the best answer below

|  |
| --- |
|  |

* **B. Cell C6 has a 50% greater chance of responding than the overall population**

Which of the following is most analagous to cloning your best customers from your marketing database?

|  |
| --- |
|  |

* **C. Spokeo Search**

Consider the attached customer database. They were recently promoted and the company accumulated their responses to the campaign as shown. They want to use this data to promote again but only to the most responsive customers. Use Excel and create a linear response model to answer the following questions. Are Age and Height meaningfully correlated for these customers?

* **D. No**

Bootstrapping is

|  |
| --- |
|  |

|  |  |
| --- | --- |
|  |  |
| **B. A method of repeatedly selecting random samples from too small a sample set for verification to find significant variables to build a model** |  |
|  |  |
|  |  |

Recency, frequency and monetary value data are classified as:

|  |
| --- |
|  |



* **C. Marketing data**