|  |  |  |  |
| --- | --- | --- | --- |
| **Name:** |  | **Date:** |  |

Thanks for your interest in our Master of Science in Data Science program. Please fill out this form to help us understand your preparation. Your answers help us, and you, to determine if you are prepared to succeed in the program.

Prerequisites for the program include a computer programming course, a statistics course, and a database course. If you have successfully completed an appropriate course, please tell me which course to look for on your transcript. If you have not taken an appropriate course, please briefly describe your experience in that area.

|  |  |
| --- | --- |
| **Topic** | **List class or describe experience:** |
| Programming (preferably object-oriented) |  |
| Statistics (probability and basic inference) |  |
| Databases (particularly SQL) |  |

Please indicate your comfort level on the scale of 0-3 (0 – Not familiar at all; 1 – Some knowledge; 2 – Good understanding; 3 – Expert) for the following subjects / topics / questions. Depending on your background, I may ask you to complete some of the tasks listed below.

|  |  |  |
| --- | --- | --- |
| **Subject** | **Topic** | **Comfort level** |
| **Programming Comfort Level** | Repetition: Write a program that prints out all integers between 10 and 30 inclusive. |  |
| Selection Statements: Write a program that reads in an integer from the user and prints out ODD if the number is odd and EVEN if the number is even. |  |
| Function/Method: Create a function/method that accepts two numbers as arguments and prints out all of the even numbers between the two numbers. Note the numbers may be sent to the function/method in any order. |  |
| Arrays/Lists: Write a program that prints out all elements in an array/list of integers that are greater than 20. |  |
| Simple Data Structures: Write a program to store basic information about zip codes. Every zip code is in exactly 1 state but a state could have many zip codes. What data structure is best for this problem? |  |
| **Programming Task:** Use any convenient programming language to implement solutions to each of the above questions. For the last one, explain which data structure you would use. | | Comfort level with this task: |
| **Statistics Comfort Level** | Descriptive Statistics: Mean, Median, Mode, Standard Deviation, Variance |  |
| Probability Rules and Concepts |  |
| Permutations and Combinations |  |
| Scatter Plot; Histogram; Correlation analysis |  |
| Probability Distributions |  |
| Simple Regression |  |
| Hypotheses Testing |  |
| Confidence Interval; Margin of Error |  |
| **Statistics Task:** You are examining if people like Coke or Pepsi equally. The p value is 0.03. Interpret this number (and explain as clearly as you can without using statistical jargon). Will your interpretation be any different if the p value is 0.06 instead of 0.03? Please explain (you may assume the level of significance for the test is 0.05). What are all the null and alternate hypotheses in this case? | | Comfort level with this task: |
| **Database / SQL Comfort Level** | SQL Basics: basics of creating tables and selecting data in various different ways. |  |
| SQL Queries: AND/OR, IN, LIKE, HAVING |  |
| Joins: inner joins, outer joins, and self joins |  |
| Modifying databases with SQL: UPDATE, DELETE, ALTER, and DROP. |  |
| **Database Task:** You have a customer table and an orders table where a customer can have many orders. Write a SQL query to return the top five customer names based on the dollar amount ordered? You can assume the tables have reasonable columns like customerFirstName, customerID, orderCost, etc. | | Comfort level with this task: |

What else would you like me to know about your preparation?

Answer only if your previous GPA was less than 3.0: This program requires that you maintain a GPA of at least 3.0 so what do you plan to do differently? Were there special circumstances before?