INTERVIEW QUESTIONS FOR DATA ANALYST

Behavioural Interview Question

Describe a time you had to learn something new to complete a task.

S: during a summer internship I was assigned a project involving Radio Frequency. T: The task was for me to create software that reads radio frequency strength from a remote control.
A: I researched some examples on the Internet and practice them until I was comfortable R: As a result, I was confident enough that I completed the project and finished it on time.

What is one current innovation in technology that excites you?

One innovation that excites me is Spark because through Spark we can manipulate billions and billions of data with as little code.

A time you avoided an obstacle

S: I was working on a team project with three other students
T: One student stopped being productive
A: I propose the team restructure the repartition of the project so we could move ahead R: As a result, we were able to finish the project on time

A time you had to request extra help to complete a project/task

S: I was working on a school project and I was stuck on some multi-threading user interface issues
T: I requested help from my pattern who had more experience with it
A: I learned enough from him and benefited from his help
R: As a result, I was able to finish the user interface within the deadline

Describe a time when you had to implement new technology in your workplace/studies.

S: During one of my internships, I had to design a storage database for recruiting

T: I worked closely with HR to gather the requirements
A: I implemented a SQL storage system to allow data attraction from resumes
R: As a result, I finished the new system on time and she gave me great feedback

When do you stop learning or have learned enough?

I know that I have learned when I'm able to move forward on a project without being stuck or looking things up.

What is your weakness?

My weakness is my attention to detail. I have had teammates that tell me that I spend a lot of time gathering requirements and specifications.

Tell us about a time you manipulated data

I worked as a data analyst for a company called Kingland System. My experience there involved writing queries for extracting data, analyzing them, and making recommendations.

Name one or two technical concepts you learned and how you applied them

I learned spark for the seek of being curious. I had a project where I had a dataset from the tweeter. Using SQL was going to be too slow so I used Spark to extract the data.

What experience do you have with extracting data? (SQL queries)

I worked as a data analyst for a company called Kingland System. My experience there involved writing queries for extracting data out of the stock market, analysing them, and making recommendations.
What is big data?

What interests you about data?

What interests me about data is the challenges that come with data analytics. Nothing is usually the same and that excites me.

What is your experience in using Excel?

My experience with Excel involves graphing data, drawing functions from input data, as well as modelling data in Excel.

<u>Case Interview Question</u>

Simple interest, compound interest, etc. For hacker Rank 2 SQL questions and 2 Python questions:

Simple interest = Principal + Principal (Rate*Time)

Compound Interest = Principal(1+Rate) ^Time

-Over Time times (years/month)

Simple profit/loss questions based on the credit card company working strategy: *case study: credit card insurance profit per customer cost*

of mailing $0.5, response rate 1%, average balance per month $1000, Insurance charges 1% of average balance monthly, customer claim rate 5% (once insured customer files a claim, the balance will be wiped and C1 lose $1000). Question: 1) Profitable? Profit per customer, 2) How to make it more profitable?

3) *What if the response rate' doubled but claims doubled?*
4) *Make a chart of the profit curve*
5) *Graph claim rate vs response rate*
6) *Graph profit vs response rate and find the response rate for maximum profit (r=1%)*

How to determine APR with Credit cards, based on data given

Simple interest = Principal + Principal (Rate*Time)

Compound Interest = Principal(1+Rate)^Time

-Over Time times (years/month)

C1 customers drop out in the middle of using their mobile app. How would you analyze and report?

I would analyse using a funnel survival guide. It's a tool that helps us see the process and steps taken by customers by giving us a visual representation of the data flow between each step. This tool will help

us to identify exactly where the customer dropped out. From there, we can develop a report with recommendations.

What are some ways you would detect credit card fraud?

- By flagging large expenses
- By flagging large expenses that are made after small expenses
- By flagging purchases that do not match the owner's usual -By flagging when the card is used in a different state or country without the card owner's call.

Originally asked: If the ratio of green to red is 2 to 7 and there are 70 total marbles, how many greens are there?
Amended question after email: If the ratio of green to red is 2 to 7 and there are 70 red marbles, how many greens are there?

Answer: Green = 2x = 0.864*2 + 7*0.864

Given the dataset, you need to perform analysis and give recommendations.

What are you gonna do if there are only 100 data for developing a model about the prediction of males or females?

Answer: It will depend on what I have available: such as First name, weight, and height.
Most likely I would go based on the world statistics prediction

Technical problems

CI: Costumer intelligence is information derived from customer data.
Missing data: It is an important element of a dataset that is missing that we need to conclude a decision on a dataset.
P-value: is the probability for a given statistical model that when the null hypothesis is true the result will be the same or more extreme than the actual results.

Doing a project about New York City taxis data. Firstly, do some basic summary statistics and box plots. The next one is to ask you to identify which trips are from/ to the three airports in NYC. Then asking you to build prediction models for tips to the driver. There are like 20 variables including continuous and categorical variables. The last question you have 4 options, I choose the data visualization one, which is to plot the pickup and drop-off locations for taxis, but also need to customize some settings to make it more informative.

Case 1:

Given APR, Interchange fee, Avg monthly balance, Avg spends every month, and loss rate of 3% calculate the profit per customer.
Now justify if it is profitable to give cash back to the customers.

Case 2:

2 **Ways of Campaigning for credit cards:**

1. **Email - 10% of applicants become customers - each representative can verify 10 email applications in an hr and is paid $25/hr**
2. **Chat - 20% of applicants become customers - each representative can respond to 4 applications in an hr and is paid $25/hr Profit per customer in both cases is $100. Which one is profitable email or chat? Draw the graph of profit vs no of applicants.**

Consider a scenario where there are only 5 representatives to handle applications.
In this case which one is more profitable email or chat? calculate the breakeven point for the number of representatives where the chat will be more profitable than email.

- An ATM makes $2 per customer for 30% of its customers. Each time a customer uses the ATM it costs 2 cents. The ATM servicing costs $3000. What is the break-even point for this ATM?

- Determine average daily cost or profit, then tweak the initial set of variables for new scenarios and finally determine a break-even point.

- The first case interview was based on the basic profit equation. profits= revenues- costs. The case was based on a credit card scenario. credit card user population is divided into two groups. for group 1 I was given all numbers and asked to calculate profit per year. For the second part, I was asked to calculate how much a different group has to spend to generate the same amount of profit. we need to make some assumptions here. I made some mistakes and the interviewer help me thru the case.

- The second case was an interview based on a campaign and calculating the benefits of the campaign. The interviewer asked me how I would calculate the profits related to the campaign. he was expecting an answer like " I would have a target and control group and based on the difference in the increase in revenue before and after the campaign I can tell what % of the increase in revenue is due to the campaign". something like that.

- Analysis of a shipping company's balance sheet

- Comparing credit card collection strategies

- Draw a line that would give Domino’s Pizza the same revenue $9 per pizza, with the y-axis (% On Time), and the x-axis ($ off for being late) - An interesting Data challenge that can be completed with any open-source tool.