 **Project** | Data Analyst Take-Home

**INTRODUCTION:** Take-home challenges are essentially mini data analysis projects. You get a business case study problem and a dataset, and then you must perform analysis, sometimes write working SQL queries, and make recommendations.

This project features one of these very take-home challenges: this one was a part of an interview for a data analyst role at Instacart.



For those unfamiliar, Instacart is an online grocery delivery and pickup service that allows users to order groceries and other household items from various local retailers through a website or mobile app. Founded in 2012, Instacart partners with a wide range of grocery stores, including national chains and local markets, to provide customers with a convenient shopping experience.

**— Directions**

And yes, this is exactly how they appeared to job candidates!

We'd love for you to analyze the data in **instacart\_data** table and share what you find. We know that you don't know much about how our team currently is run, but that's okay. This data set includes information on orders, order location, customer ratings, and any issues reported by the customer for a set of orders.

Please analyze the data and share with us …

1. … any observations about our business.
2. How would you staff the Customer Support Team?

Please compile your analysis into a presentation to convey your findings. Use the dataset as necessary to substantiate your claims.

**— Rubric**

In the real world, you often won't have a rubric to guide your take-home assessments. However, for this exercise, we're providing one to help you understand how your work will be evaluated and to give you detailed feedback. Use this as a tool to refine your approach and improve your skills!

**— Presentation**

* Create a slide presentation (e.g., PowerPoint, Google Slides) with the results of the tasks, emphasizing the interpretation of your findings.
  + Take a look at [SlidesGo](https://slidesgo.com/) and [Slides Carnival](https://www.slidescarnival.com/) for fantastic, free slide templates!
* Submit your presentation as a PDF.
* You will prepare a 5-10 minute presentation and submit a recording of your presentation. This practice is valuable for simulating live presentations, a crucial skill in data analytics.

Have fun, and good luck!

**— Scoring**

* **Exploratory Data Analysis (EDA):** /5
* **Business Observations:** /5
* **Staffing Recommendations for Customer Support Team:** /5
* **Presentation and Communication:** /5

**Total Score:** /20

**Sample Queries:**

**Orders by Region**

**select**

**region,**

**count(order\_id)**

**from instacart\_data**

**group by 1**

|  |  |
| --- | --- |
| **chi** | **6430** |
| **nyc** | **1290** |
| **sf** | **7237** |
|  |  |

**- Get orders by region and month**

**select**

**region,**

**strftime('%m', order\_date) AS month,**

**count(order\_id)**

**from instacart\_data**

**group by 1,2**

|  |  |  |
| --- | --- | --- |
| **chi** | **05** | **6055** |
| **chi** | **06** | **375** |
| **nyc** | **05** | **1246** |
| **nyc** | **06** | **44** |
| **sf** | **05** | **6818** |
| **sf** | **06** | **419** |

**-- Which region has the least problem with order issues?**

**with order\_count\_by\_issue as (**

**select**

**region,**

**-- strftime('%m', order\_date) AS month,**

**type\_of\_issue\_reported,**

**count(order\_id) as count\_of\_issue,**

**sum(count(order\_id))over(partition by region) as region\_total**

**from instacart\_data**

**group by 1,2**

**order by 1, 3 desc**

**)**

**select**

**region,**

**type\_of\_issue\_reported,**

**round((count\_of\_issue \* 1.0)/ region\_total ,3) \*100 as pct\_of\_total**

**from order\_count\_by\_issue**

**where type\_of\_issue\_reported = 'No issue reported'**

|  |  |  |
| --- | --- | --- |
| **chi** | **No issue reported** | **95.8** |
| **nyc** | **No issue reported** | **90.5** |
| **sf** | **No issue reported** | **90.4** |

**Chicago has the best percentage.**