ROBLOX DATA SCIENTIST INTERVIEW

- **SQL Coding Challenge**
  - Write a SQL query to find the total number of users who have played Roblox in the last 30 days.
  - Write a SQL query to find the average playtime for each game in Roblox.
  - Write a SQL query to find the total number of purchases made in the last 90 days.
  - Write a SQL query to find the top 5 games in Roblox based on the number of players.
  - Write an SQL query to find the top 10 most played Roblox games in the last 30 days, along with their average playtime per user.
  - Given a table with user purchase data, write a SQL query to identify users who haven't purchased anything in the last 6 months and might churn, along with their key demographics and engagement metrics.
- **Python** (pandas, numpy), **Probability & Statistics, Hypothesis Testing, A/B Testing**
  - Identify key metrics, trends, and analytical needs with an emphasis on improving and scaling systems
  - P-value, confidence interval
  - https://www.interviewquery.com/questions/three-zebras
  - https://www.interviewquery.com/questions/500-cards
  - https://www.interviewquery.com/questions/bootstrapping-confidence-intervals
  - https://www.interviewquery.com/questions/replace-words-with-stems
  - https://www.interviewquery.com/questions/estimated-rounds
  - https://www.interviewquery.com/questions/find-bigrams
  - https://www.interviewquery.com/questions/statistically-significant-test
  - https://www.interviewquery.com/questions/new-ui-effect
  - Number of Trusted Contacts of a Customer - LeetCode
  - Roblox - LeetCode
  - Design Browser History - LeetCode
  - You are given an array of positive integers numbers and a positive integer separation. Your task is to find the minimal difference between elements of numbers where their indices are at least separation apart.
    Task: Formally, you are tasked to find:
    min(|numbers[i]−numbers[j]|)
    where |i−j| ≥ separation.

    Solution: Time Complexity: O(n), Space Complexity: O(1)

```python
def solution(numbers, separation):
    # Initialize the minimum difference with a large number
    min_diff = float('inf')

    # Loop through the array to find pairs of numbers that meet the separation requirement
    for i in range(len(numbers) - separation):
        # Calculate the difference between the current element and the element at the requ
        diff = abs(numbers[i] - numbers[i + separation])

        # Update the minimum difference if a new smaller difference is found
        if diff < min_diff:
            min_diff = diff

    return min_diff

# Example usage:
numbers = [1, 5, 4, 10, 9]
separation = 3
print(solution(numbers, separation))  # Output: 4

numbers = [3, 10, 5, 8]
separation = 1
print(solution(numbers, separation))  # Output: 2
```

- **ML Interview: Supervised & Unsupervised learning, ML Design**
  - Code K-Means
  - What approaches do you use to assess model performance?
  - How would you build a recommendation system for Roblox games based on user preferences and engagement data? What algorithms and techniques would you consider?
  - How would you visualize user engagement metrics on the Roblox platform using Python libraries like Pandas and Matplotlib?
  - How would you develop a strategy for anomaly detection in Roblox game performance data?
  - How would you build a collaborative filtering recommendation system for Roblox games, considering the dynamic nature of user preferences and content creation?
- **Basics of Data pipelines: Dockers, Kubernetes, Airflow, Snowflake, Big data**
- **Online Assessment Games**: Series of online games (2-3) to test cognitive skills
  - selecting infected animals based on a map of infected foxes at different time.
  - Building as many as possible cars to reach the destination through multiple barriers. You can add 3 accessories (overall there are 7-9) to the car and make it overcome the obstacles in the path.
- **Product case study questions (check specially gaming industry)**
  - product case questions about their Avatar shop, and how to test my recommendations.
  - A/B testing with Avatar shop.
  - Evaluate the impact of listing banner ads on Roblox.
  - Discuss any strategies you used to optimize your performance in the online games?
  - Discuss the importance of data governance and security at Roblox and how you would contribute to ensuring responsible and ethical data practices within the company.
  - Understanding of Roblox platform, including player behavior patterns, engagement strategies, monetization tactics, and community building
- **Behavioral questions and past projects/experiences.**

<mark>**Jin Guo 1st Discussion**</mark>

- Can you Tell me about the most complex project you have ever completed including choices and technologies in layman's terms?
- What have you done to improve the effectiveness (e.g. accuracy, reproducibility or scalability) of data analytics work at your company?
- What is a margin of error?
- Can you give an example of selection bias.
- What are the assumptions required for a Linear Regression?