**NetID: jc12300**

**Name: Jie Cheng**

**Course: CSGY-6083-Principles of Database Systems**

**Section: B**

**Date of submission: Nov 25, 2023**

****

CS-GY 6083 - B, Fall 2023

Principles of Database Systems

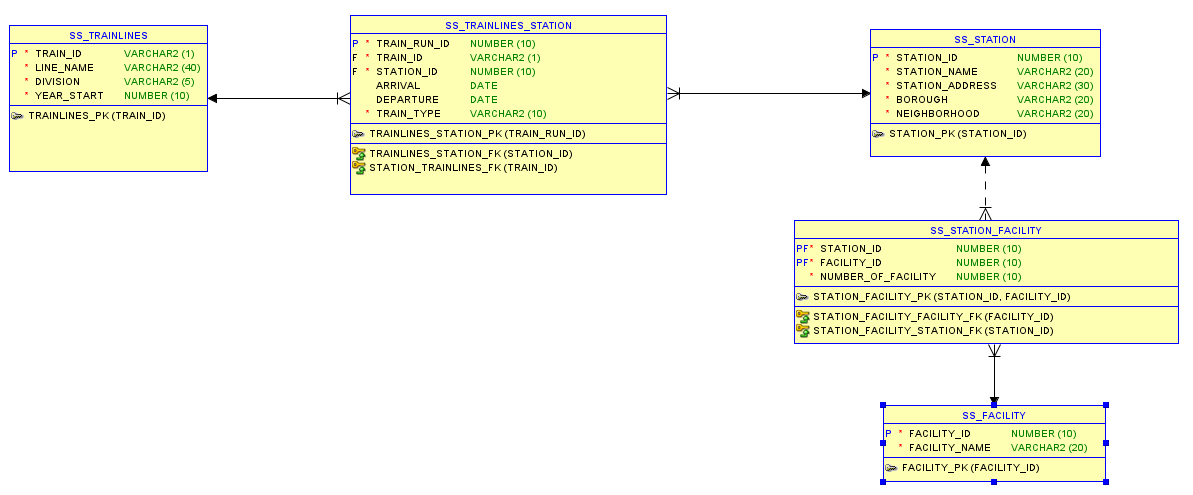
ASSIGNMENT: 3 [100 Points, 5% weight to the Final grades]

Please submit your assignment on NYU Brightspace in a single PDF document attachment. Please mention Student ID, Name, Course, Section Number, and date of submission on first page of your submission. Please use SQL file (DDL and DML code) for writing SQLs in Oracle. You may create equivalent DDL and DML code for MySQL and write SQLs in MySQL. All the tables should be created with your initial as a prefix. DO NOT change data in tables. You can write SQL in either Oracle or MySQL.

**Problem A: 50 points [each question has 10 points]**

# The New York City Subway is a rapid transit system in the New York City boroughs of Manhattan, Brooklyn, Queens, and the Bronx. It is owned by the government of New York City and leased to the New York City Transit Authority, an affiliate agency of the state-run Metropolitan Transportation Authority (MTA). Opened on October 27, 1904, the New York City Subway is one of the world's oldest public transit systems, one of the most used, and the one with the most stations, with 472 stations in operation and 36 train lines.

SAMTA (Subway Administration of Metropolitan Transportation Authority) is the NYU affiliated data analytics startup company. SAMTA has undertaken a database project to normalize subway data systems. For relational schema model in figure, write SQL statements to answer following questions. You don’t need to create a relational model.



1. **List all the details of the station which has the highest number of facilities in Elevator.**

文本

描述已自动生成



1. **List details of each train line and their highest run time. Your result should have TRAIN\_ID, LINE\_NAME, ARRIVVAL, DEPARTURE,**

文本

描述已自动生成

图形用户界面, 文本

描述已自动生成

1. **Find the top two boroughs in terms of numbers of station facilities. Your result should have the name of the borough and number of facilities.**

文本

描述已自动生成

图形用户界面, 应用程序

描述已自动生成

1. **List station name, borough, neighborhood, facility name, and number of facilities. Arrange the result in descending order of number of facilities.**

文本

描述已自动生成

文本, 日历

描述已自动生成

1. **List the train line that has maximum number of stations.**

文本

描述已自动生成

图形用户界面, 应用程序

中度可信度描述已自动生成

**Submit: For each of questions submit a) SQL b) SQL result. All SQL and corresponding results must be visible clearly on screenshots.**

**Problem 2: 20 points [each question has 10 points]**

 A yellow sign with black text

Description automatically generated

 A yellow sign with black text

Description automatically generated

 A purple rectangle with white text

Description automatically generated

 A purple rectangle with white text

Description automatically generated

**Submission:**

**Create tables as above with your initial as prefix such as AP\_FLIGHT, AP\_SIZE. Populate the same data as shown in pictures. For Q1 and Q2 write the SQLs that produce the desired results as Output 1 and Output 2. Submit screenshots of your SQL queries and their respective results. You can use Oracle or MySQL relational database. All SQL and corresponding results must be visible clearly on screenshots.**

**Q1**

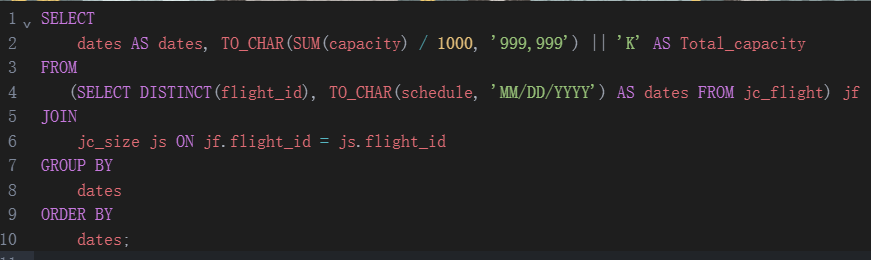
手机屏幕的截图

描述已自动生成

图形用户界面, 应用程序

描述已自动生成

**Q2**

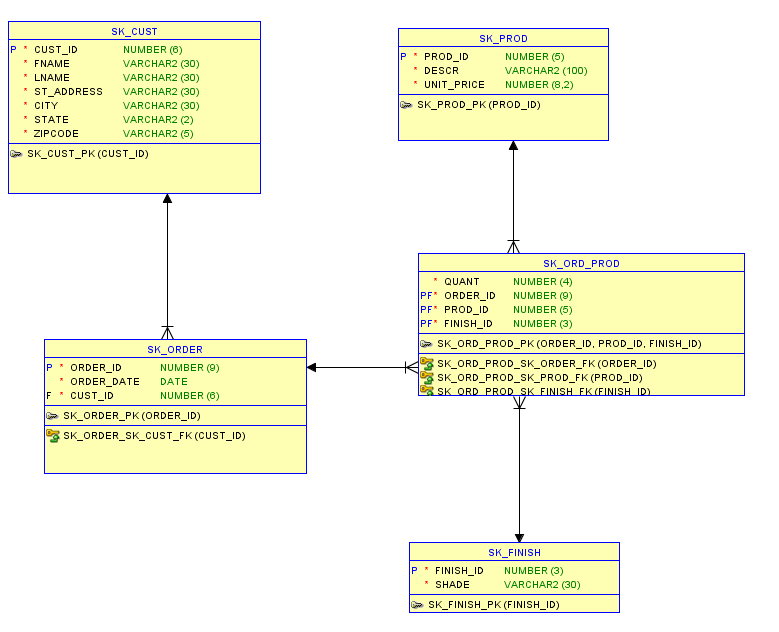


图形用户界面, 应用程序

描述已自动生成

**Problem 3: 30 points**

**For a given relational model below, please find attached file containing DDL and DMLs. You may create equivalent code for MySQL. Create tables and insert data by replacing SK with your own initial. You can do this assignment either in Oracle or MySQL. You don’t need to draw logical/relational models.**



1. **For this relational model of a furniture company, create a read-only database view that represents the following dataset. Customer ID, Customer Name (both First and Last name), Order\_Id, order date, each product in order with description, quantity, Unit\_Price, Total price of each product, and Finish shade. Sort the dataset in order of total order amount. Give appropriate column names in view. Restrict dataset to represent only those orders which have total value over $1000. Once view is created submit View code and then retrieve result of view using SQL query.**

文本

描述已自动生成

屏幕上有字

描述已自动生成

1. **Find top 3 products and their finish shade in terms of total quantities sold between 01/01/2018 and 12/31/2021. Your result dataset should have Product\_Id, Product Description, shade, and total quantity sold.**

文本

描述已自动生成

表格

描述已自动生成

1. **Find products and their finish shade that have not been sold in during the months of Oct, Nov, and Dec in 2020**

文本

描述已自动生成

图形用户界面, 文本, 日历

描述已自动生成

**Submission:**

**For each question write an SQL query and submit both SQL and a clearly visible screenshot of corresponding result underneath each question. Make appropriate use of column alias and built in functions in your SQL queries.**