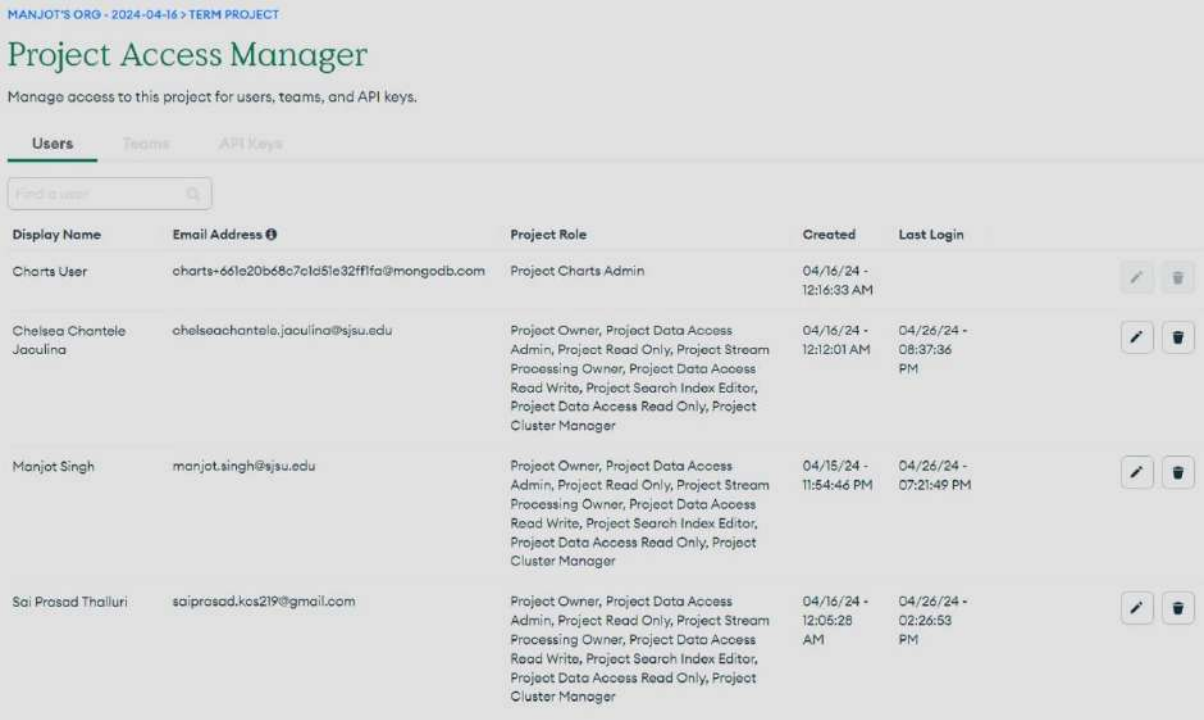
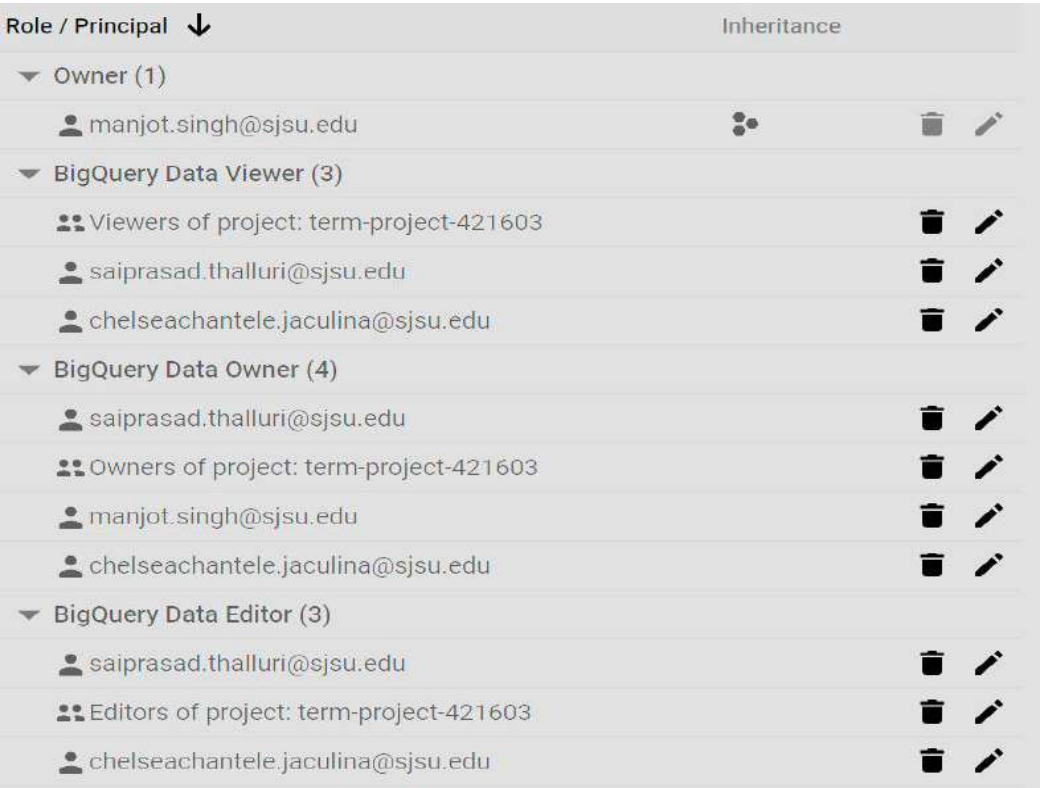
DATA 225 Database Project Criteria Evidences

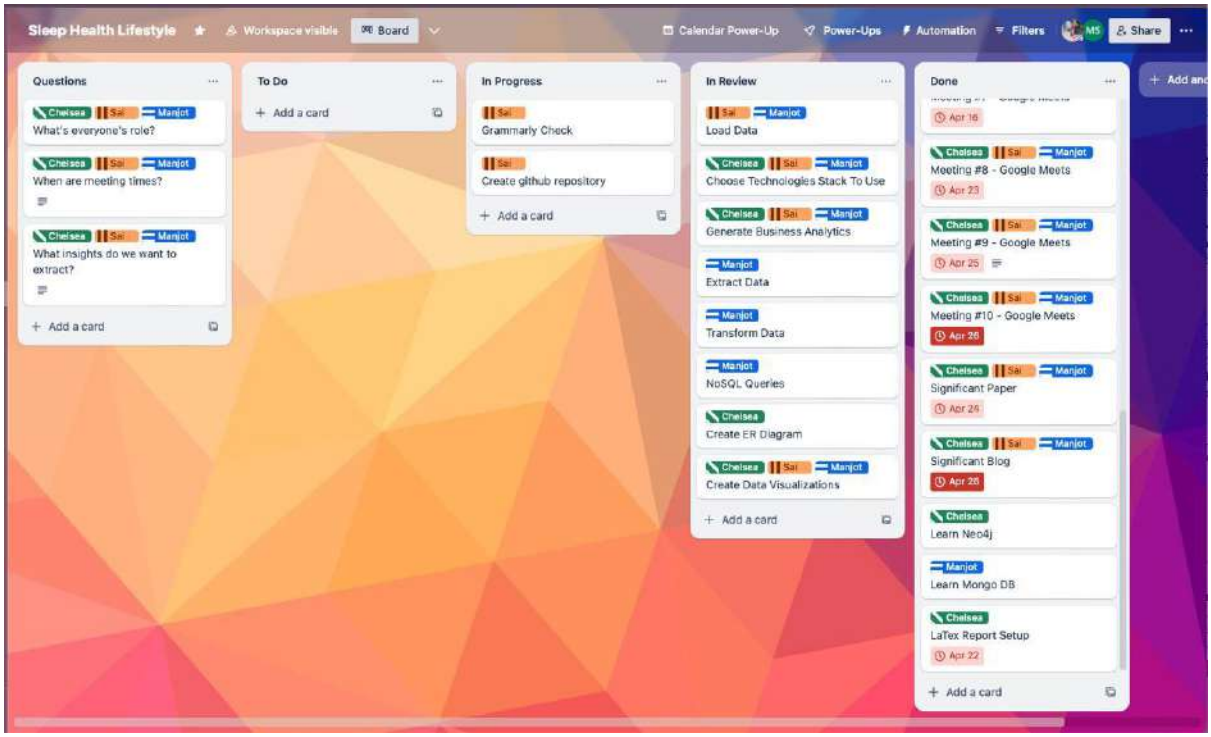
Practiced pair programming?

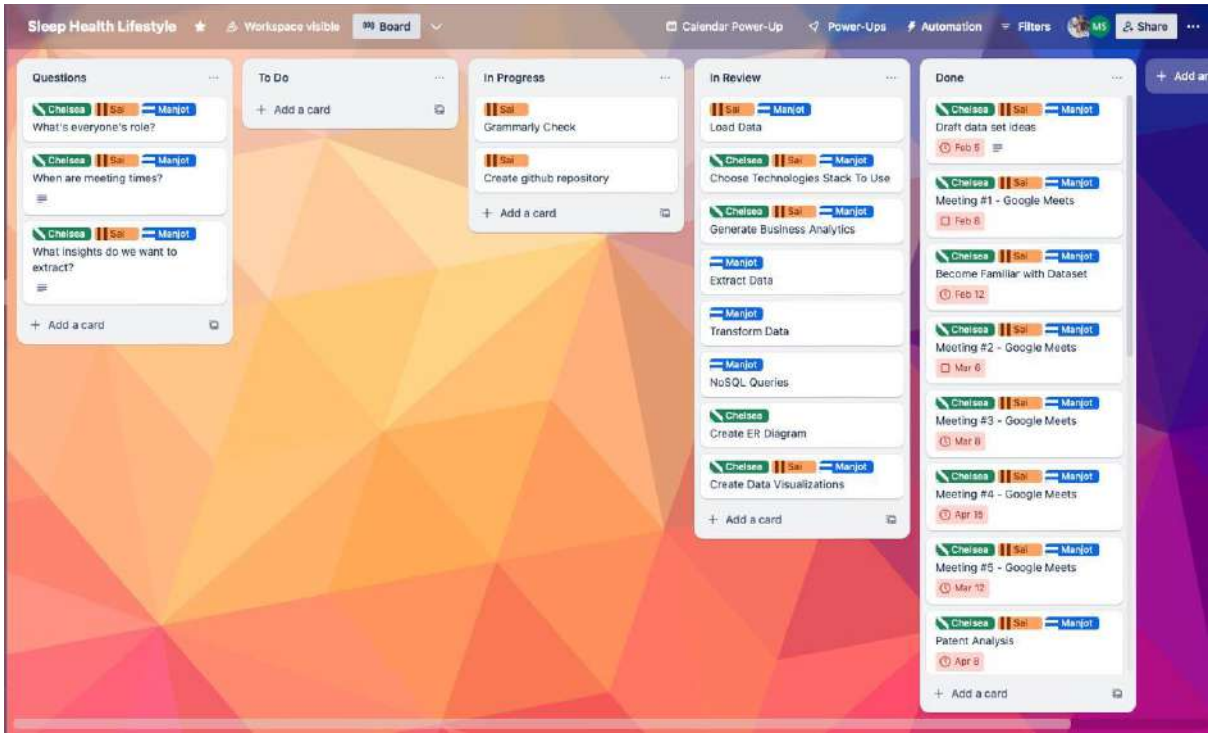




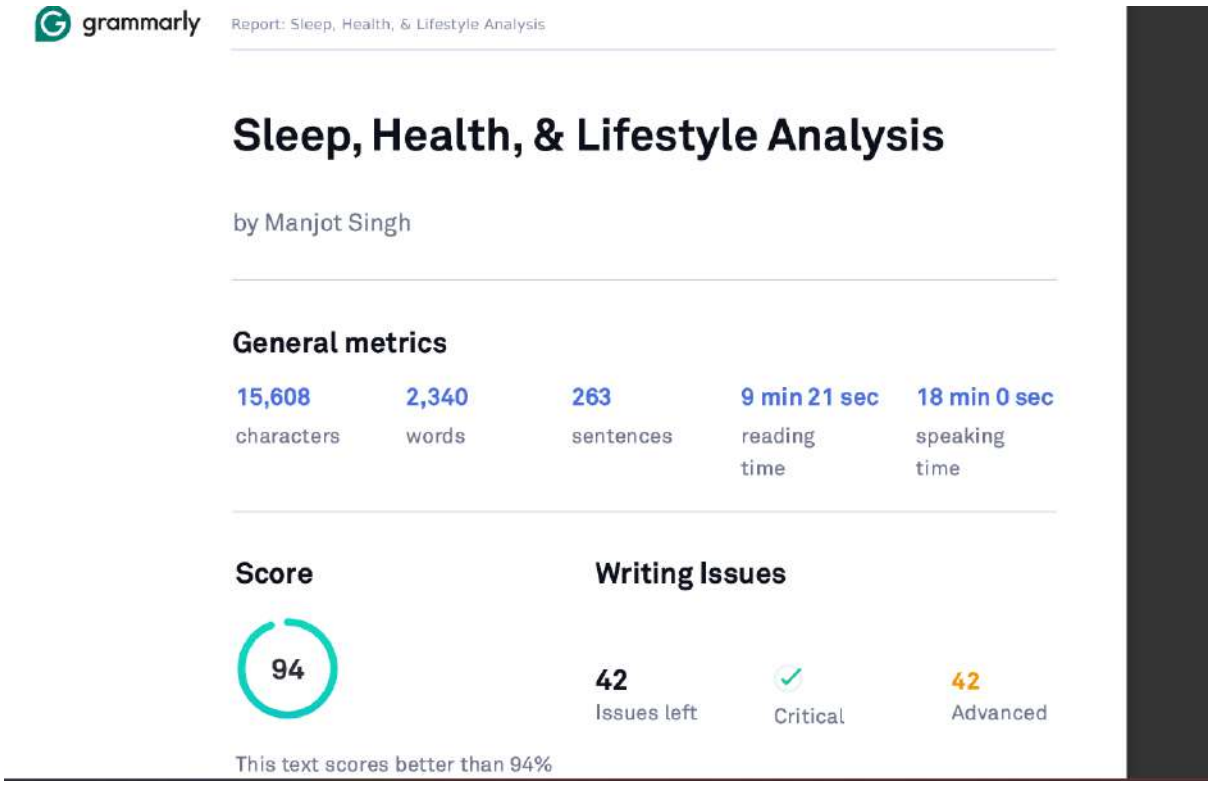
Practiced agile / scrum (1-week sprints)?

Trello:

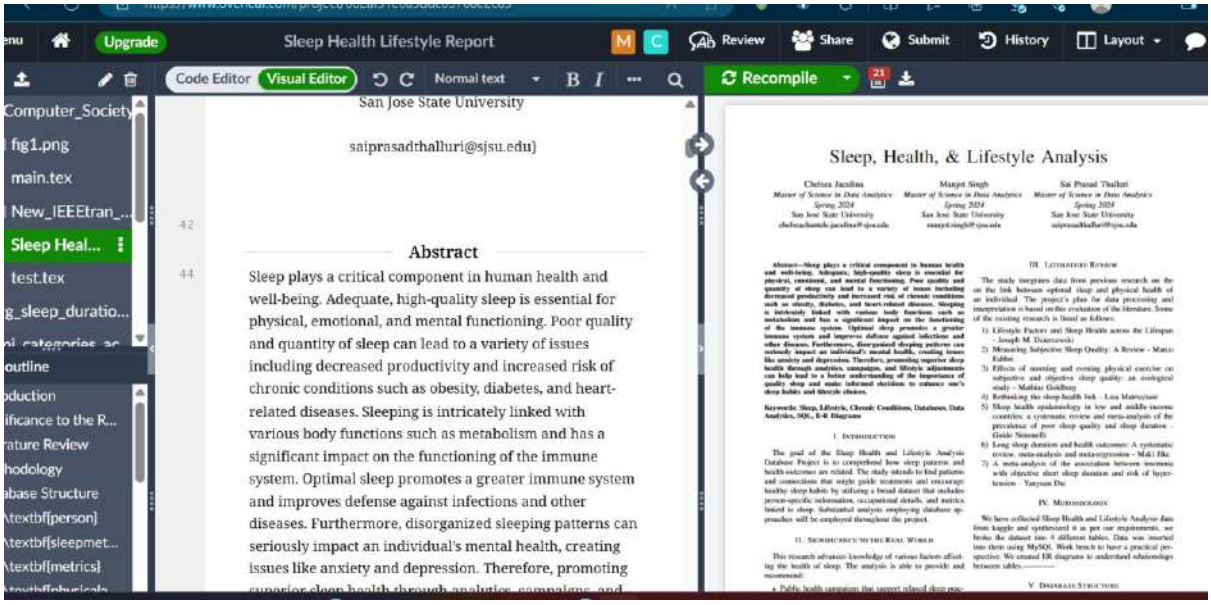




Used Grammarly / other tools for language?



Used unique tools: LATEX



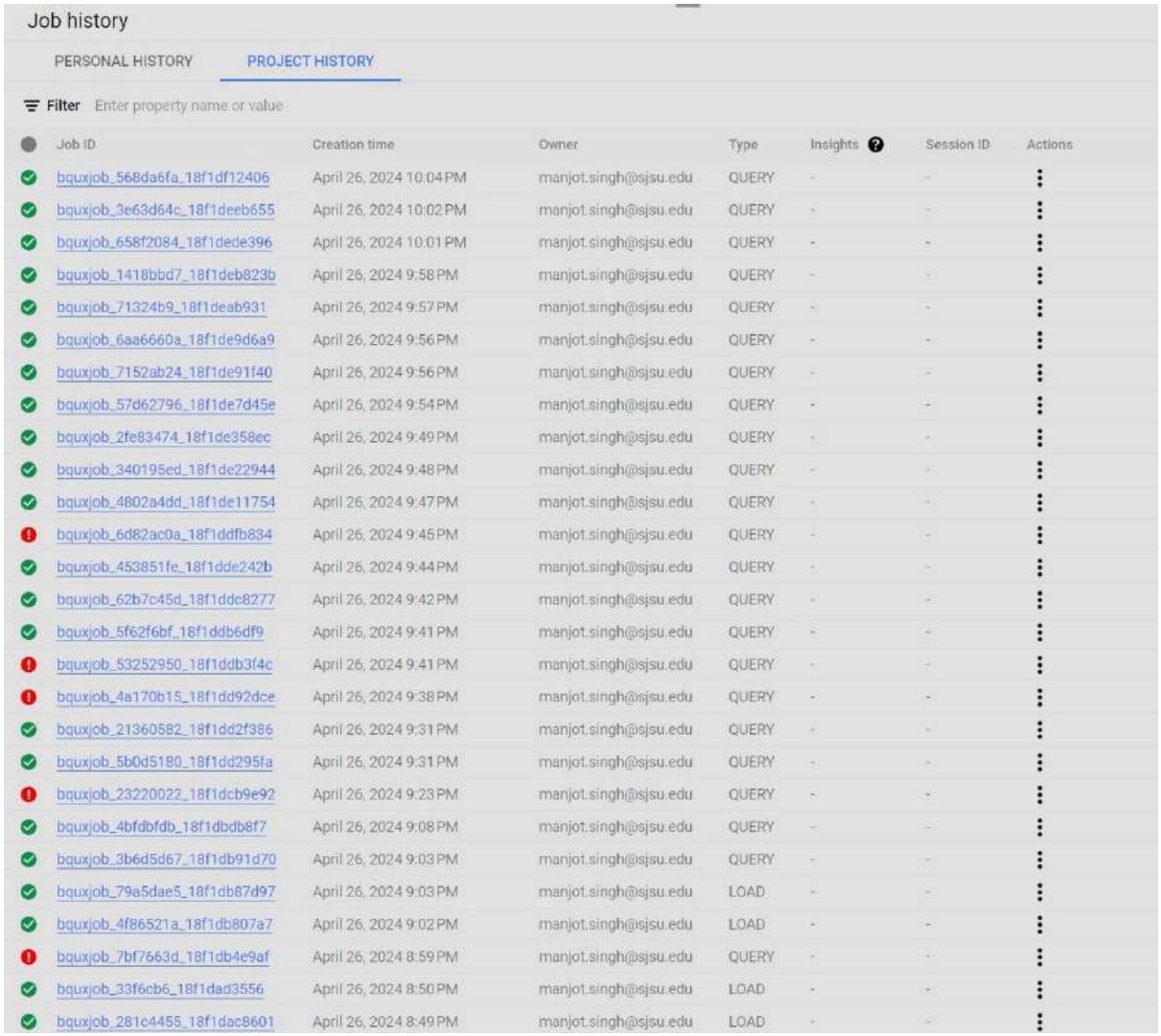
[Sleep Health Lifestyle Report - Online LaTeX Editor Overleaf]

Used a new database or data warehouse tool not covered in the HW or class

Google Big Query Warehouse Tool

```sql
CREATE OR REPLACE TABLE Sleep_Health.TransformedData AS
SELECT
  P.PersonID,
  P.Gender,
  P.Age,
  P.Occupation,
  SM.SleepDuration,
  SM.QualityofSleep,
  PA.PhysicalActivityLevel,
  CASE
    WHEN M.StressLevel <= 3 THEN 'Low Stress'
    WHEN M.StressLevel > 3 AND M.StressLevel <= 6 THEN 'Moderate Stress'
    ELSE 'High Stress'
  END AS StressLevelCategory,
  M.BMICategory,
  M.BloodPressure,
  M.HeartRate,
  PA.DailySteps,
  SM.SleepDisorder
FROM
  Sleep_Health.Person AS P
JOIN
  Sleep_Health.Sleep_Details AS SM ON P.PersonID = SM.PersonID
JOIN
  Sleep_Health.Physical_Activity AS PA ON P.PersonID = PA.PersonID
JOIN
  Sleep_Health.Metrics AS M ON P.PersonID = M.PersonID;
```

Row	PersonID	Gender	Age	Occupation	SleepDuration	QualityofSleep	PhysicalActivityl	StressLevelCategory
1	18	Male	29	Doctor	6.0	6	30	High Stress
2	14	Male	29	Doctor	6.0	6	30	High Stress
3	16	Male	29	Doctor	6.0	6	30	High Stress
4	15	Male	29	Doctor	6.0	6	30	High Stress
5	62	Male	32	Doctor	6.0	6	30	High Stress
6	58	Male	32	Doctor	6.0	6	30	High Stress
7	55	Male	32	Doctor	6.0	6	30	High Stress
8	53	Male	32	Doctor	6.0	6	30	High Stress
9	59	Male	32	Doctor	6.0	6	30	High Stress
10	56	Male	32	Doctor	6.0	6	30	High Stress
11	61	Male	32	Doctor	6.0	6	30	High Stress
12	77	Male	33	Doctor	6.0	6	30	High Stress
13	80	Male	33	Doctor	6.0	6	30	High Stress
14	76	Male	33	Doctor	6.0	6	30	High Stress
15	68	Male	33	Doctor	6.0	6	30	High Stress
16	75	Male	33	Doctor	6.0	6	30	High Stress
17	78	Male	33	Doctor	6.0	6	30	High Stress
18	79	Male	33	Doctor	6.0	6	30	High Stress
19	269	Female	49	Nurse	6.0	6	90	High Stress



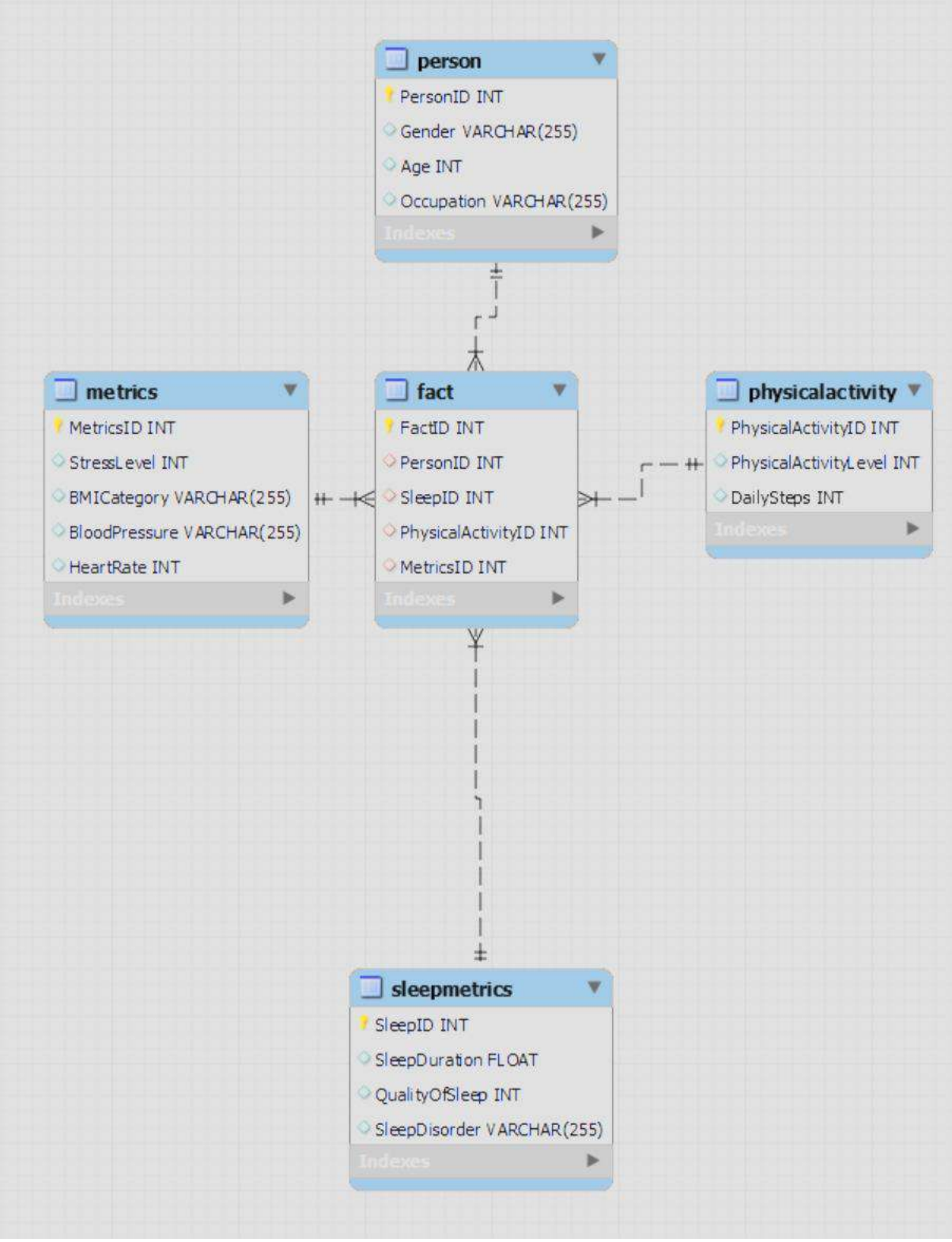
Job history

PERSONAL HISTORY | PROJECT HISTORY

Filter Enter property name or value

Job ID	Creation time	Owner	Type	Insights	Session ID	Actions
bquxjob_568da6fa_18f1df12406	April 26, 2024 10:04 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_3e63d64c_18f1deeb655	April 26, 2024 10:02 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_658f2084_18f1dede396	April 26, 2024 10:01 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_1418bbd7_18f1deb823b	April 26, 2024 9:58 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_71324b9_18f1deab931	April 26, 2024 9:57 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_6aa6660a_18f1de9d6a9	April 26, 2024 9:56 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_7152ab24_18f1de91f40	April 26, 2024 9:56 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_57d62796_18f1de7d45e	April 26, 2024 9:54 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_2fe83474_18f1de358ec	April 26, 2024 9:49 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_340195ed_18f1de22944	April 26, 2024 9:48 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4802a4dd_18f1de11754	April 26, 2024 9:47 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_6d82ac0a_18f1ddfb834	April 26, 2024 9:45 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_453851fe_18f1dde242b	April 26, 2024 9:44 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_62b7c45d_18f1ddc8277	April 26, 2024 9:42 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_5f62f6bf_18f1ddb6df9	April 26, 2024 9:41 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_53252950_18f1ddb3f4c	April 26, 2024 9:41 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4a170b15_18f1dd92dce	April 26, 2024 9:38 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_21360582_18f1dd2f386	April 26, 2024 9:31 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_5b0d5180_18f1dd295fa	April 26, 2024 9:31 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_23220022_18f1dcb9e92	April 26, 2024 9:23 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4bfdbfdb_18f1dbdb8f7	April 26, 2024 9:08 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_3b6d5d67_18f1db91d70	April 26, 2024 9:03 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_79a5dae5_18f1db87d97	April 26, 2024 9:03 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_4f86521a_18f1db807a7	April 26, 2024 9:02 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_7bf7663d_18f1db4e9af	April 26, 2024 8:59 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_33f6cb6_18f1dad3556	April 26, 2024 8:50 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_281c4455_18f1dac8601	April 26, 2024 8:49 PM	manjot.singh@sjsu.edu	LOAD	-	-	

Used appropriate data modeling techniques



Used ETL tool

Google Big Query

EXTRACT

```sql
SELECT
  P.PersonID,
  P.Gender,
  P.Age,
  SM.QualityofSleep,
  PA.PhysicalActivityLevel,
  M.StressLevel,
  M.BMICategory,
  M.BloodPressure,
  M.HeartRate,
  PA.DailySteps,
FROM
  Sleep_Health.Person AS P
JOIN
  Sleep_Health.Sleep_Details AS SM ON P.PersonID = SM.PersonID
JOIN
  Sleep_Health.Physical_Activity AS PA ON P.PersonID = PA.PersonID
JOIN
  Sleep_Health.Metrics AS M ON P.PersonID = M.PersonID
WHERE
  P.Age >= 30
  AND SM.QualityofSleep >= 7
  AND PA.PhysicalActivityLevel >= 3
ORDER BY
  P.PersonID;
```

Press Alt+F1 for Accessibility Optio

Query results

SAVE RESULTS | EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	PersonID	Gender	Age	QualityofSleep	PhysicalActivityLeve	StressLevel	BMICategory	BloodPressure
1	20	Male	30	7	75	6	Normal	120/80
2	21	Male	30	7	75	6	Normal	120/80
3	22	Male	30	7	75	6	Normal	120/80
4	23	Male	30	7	75	6	Normal	120/80
5	24	Male	30	7	75	6	Normal	120/80
6	25	Male	30	7	75	6	Normal	120/80

TRANSFORM

```sql
CREATE OR REPLACE TABLE Sleep_Health.TransformedData AS
SELECT
  P.PersonID,
  P.Gender,
  P.Age,
  P.Occupation,
  SM.SleepDuration,
  SM.QualityofSleep,
  PA.PhysicalActivityLevel,
  CASE
    WHEN M.StressLevel <= 3 THEN 'Low Stress'
    WHEN M.StressLevel > 3 AND M.StressLevel <= 6 THEN 'Moderate Stress'
    ELSE 'High Stress'
  END AS StressLevelCategory,
  M.BMICategory,
  M.BloodPressure,
  M.HeartRate,
  PA.DailySteps,
  SM.SleepDisorder
FROM
  Sleep_Health.Person AS P
JOIN
  Sleep_Health.Sleep_Details AS SM ON P.PersonID = SM.PersonID
JOIN
  Sleep_Health.Physical_Activity AS PA ON P.PersonID = PA.PersonID
JOIN
  Sleep_Health.Metrics AS M ON P.PersonID = M.PersonID;
```

Row	PersonID	Gender	Age	Occupation	SleepDuration	QualityofSleep	PhysicalActivityl	StressLevelCategory
1	18	Male	29	Doctor	6.0	6	30	High Stress
2	14	Male	29	Doctor	6.0	6	30	High Stress
3	16	Male	29	Doctor	6.0	6	30	High Stress
4	15	Male	29	Doctor	6.0	6	30	High Stress
5	62	Male	32	Doctor	6.0	6	30	High Stress
6	58	Male	32	Doctor	6.0	6	30	High Stress
7	55	Male	32	Doctor	6.0	6	30	High Stress
8	53	Male	32	Doctor	6.0	6	30	High Stress
9	59	Male	32	Doctor	6.0	6	30	High Stress
10	56	Male	32	Doctor	6.0	6	30	High Stress
11	61	Male	32	Doctor	6.0	6	30	High Stress
12	77	Male	33	Doctor	6.0	6	30	High Stress
13	80	Male	33	Doctor	6.0	6	30	High Stress
14	76	Male	33	Doctor	6.0	6	30	High Stress
15	68	Male	33	Doctor	6.0	6	30	High Stress
16	75	Male	33	Doctor	6.0	6	30	High Stress
17	78	Male	33	Doctor	6.0	6	30	High Stress
18	79	Male	33	Doctor	6.0	6	30	High Stress
19	269	Female	49	Nurse	6.0	6	90	High Stress

Load

```sql
SELECT
  f.FactID,
  f.PersonID,
  f.SleepID,
  f.PhysicalActivityID,
  f.MetricsID,
  p.Gender,
  p.Age,
  p.Occupation,
  s.SleepDuration,
  s.QualityOfSleep,
  s.SleepDisorder,
  pa.PhysicalActivityLevel,
  pa.DailySteps,
  m.StressLevel,
  m.BMICategory,
  m.BloodPressure,
  m.HeartRate
FROM
  Sleep_Health.Fact_Table f
JOIN
  Sleep_Health.Person p ON f.PersonID = p.PersonID
JOIN
  Sleep_Health.Sleep_Details s ON f.SleepID = s.SleepID
JOIN
  Sleep_Health.Physical_Activity pa ON f.PhysicalActivityID = pa.PhysicalActivityID
JOIN
  Sleep_Health.Metrics m ON f.MetricsID = m.MetricsID
```

Press Alt+F1 for Accessibility O

Query results

SAVE RESULTS | EXPLORE DATA

JOB INFORMATION | RESULTS | CHART | JSON | EXECUTION DETAILS | EXECUTION GRAPH

Row	FactID	PersonID	SleepID	PhysicalActivityID	MetricsID	Gender	Age	Occupation
1	1	1	1	1	1	Male	27	Software Engineer
2	2	2	2	2	2	Male	28	Doctor
3	3	3	3	3	3	Male	28	Doctor
4	4	4	4	4	4	Male	28	Sales Representative
5	5	5	5	5	5	Male	28	Sales Representative
6	6	6	6	6	6	Male	28	Software Engineer

Used RDBMS

Mysql:

```sql
SELECT
    CASE
        WHEN Age < 20 THEN 'Under 20'
        WHEN Age >= 20 AND Age < 30 THEN '20-29'
        WHEN Age >= 30 AND Age < 40 THEN '30-39'
        WHEN Age >= 40 AND Age < 50 THEN '40-49'
        ELSE '50+'
    END AS AgeGroup,
    SUM(CASE WHEN SleepDisorder = 'Insomnia' THEN 1 ELSE 0 END) AS Insomnia,
    SUM(CASE WHEN SleepDisorder = 'Sleep Apnea' THEN 1 ELSE 0 END) AS SleepApnea,
    SUM(CASE WHEN SleepDisorder = 'None' THEN 1 ELSE 0 END) AS None
FROM Person AS P
JOIN Sleepmetrics AS SM ON P.PersonID = SM.SleepID
GROUP BY AgeGroup;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

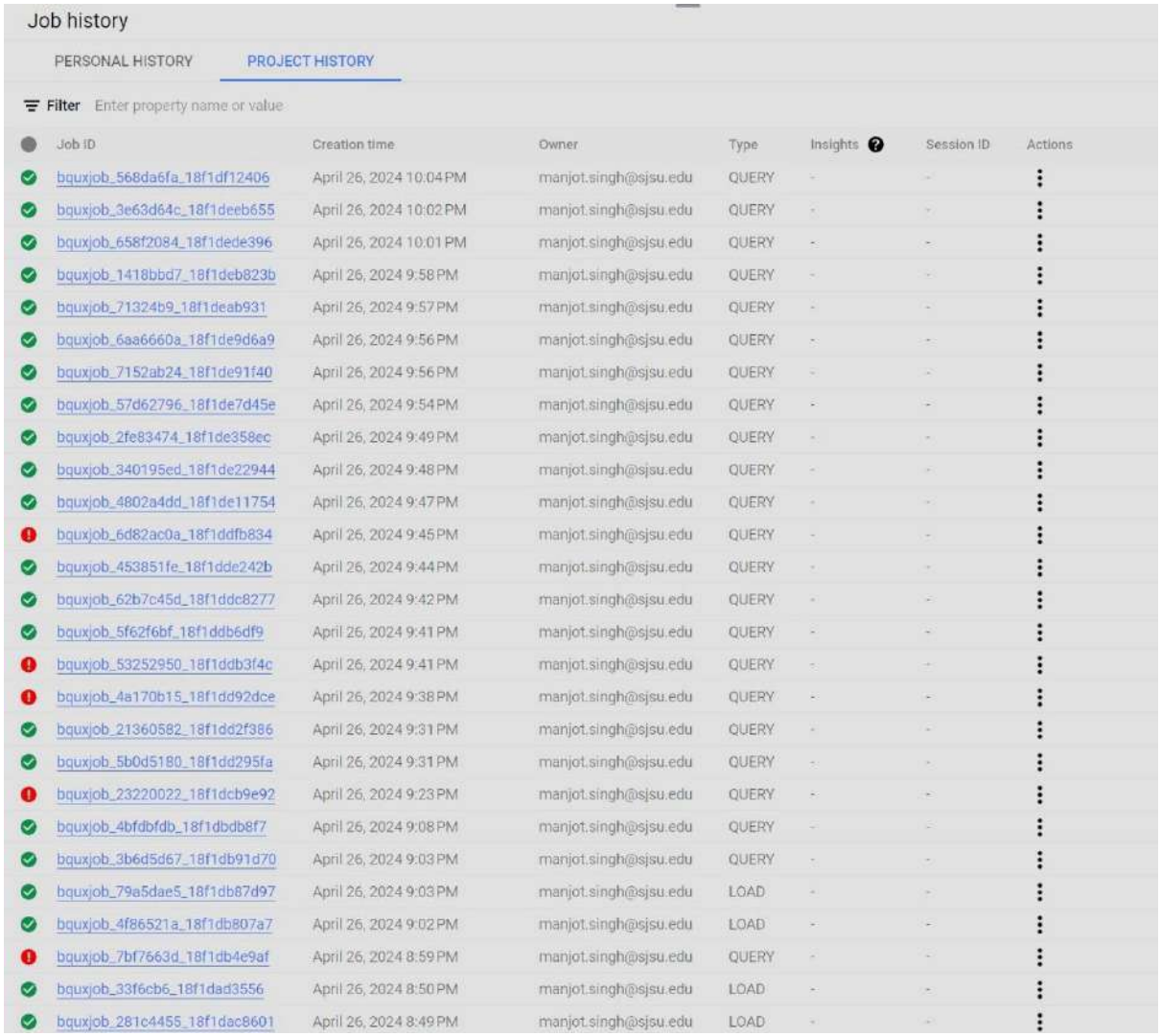
AgeGroup	Insomnia	SleepApnea	None
20-29	3	4	12
30-39	6	9	127
40-49	60	15	42
50+	8	50	38

Used Datawarehouse

Google Big Query:

```sql
CREATE OR REPLACE TABLE Sleep_Health.TransformedData AS
SELECT
  P.PersonID,
  P.Gender,
  P.Age,
  P.Occupation,
  SM.SleepDuration,
  SM.QualityofSleep,
  PA.PhysicalActivityLevel,
  CASE
    WHEN M.StressLevel <= 3 THEN 'Low Stress'
    WHEN M.StressLevel > 3 AND M.StressLevel <= 6 THEN 'Moderate Stress'
    ELSE 'High Stress'
  END AS StressLevelCategory,
  M.BMICategory,
  M.BloodPressure,
  M.HeartRate,
  PA.DailySteps,
  SM.SleepDisorder
FROM
  Sleep_Health.Person AS P
JOIN
  Sleep_Health.Sleep_Details AS SM ON P.PersonID = SM.PersonID
JOIN
  Sleep_Health.Physical_Activity AS PA ON P.PersonID = PA.PersonID
JOIN
  Sleep_Health.Metrics AS M ON P.PersonID = M.PersonID;
```

Row	PersonID	Gender	Age	Occupation	SleepDuration	QualityofSleep	PhysicalActivityl	StressLevelCategory
1	18	Male	29	Doctor	6.0	6	30	High Stress
2	14	Male	29	Doctor	6.0	6	30	High Stress
3	16	Male	29	Doctor	6.0	6	30	High Stress
4	15	Male	29	Doctor	6.0	6	30	High Stress
5	62	Male	32	Doctor	6.0	6	30	High Stress
6	58	Male	32	Doctor	6.0	6	30	High Stress
7	55	Male	32	Doctor	6.0	6	30	High Stress
8	53	Male	32	Doctor	6.0	6	30	High Stress
9	59	Male	32	Doctor	6.0	6	30	High Stress
10	56	Male	32	Doctor	6.0	6	30	High Stress
11	61	Male	32	Doctor	6.0	6	30	High Stress
12	77	Male	33	Doctor	6.0	6	30	High Stress
13	80	Male	33	Doctor	6.0	6	30	High Stress
14	76	Male	33	Doctor	6.0	6	30	High Stress
15	68	Male	33	Doctor	6.0	6	30	High Stress
16	75	Male	33	Doctor	6.0	6	30	High Stress
17	78	Male	33	Doctor	6.0	6	30	High Stress
18	79	Male	33	Doctor	6.0	6	30	High Stress
19	269	Female	49	Nurse	6.0	6	90	High Stress



Job history

PERSONAL HISTORY | PROJECT HISTORY

Filter Enter property name or value

Job ID	Creation time	Owner	Type	Insights	Session ID	Actions
bquxjob_568da6fa_18f1df12406	April 26, 2024 10:04 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_3e63d64c_18f1deeb655	April 26, 2024 10:02 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_658f2084_18f1dede396	April 26, 2024 10:01 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_1418bbd7_18f1deb823b	April 26, 2024 9:58 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_71324b9_18f1deab931	April 26, 2024 9:57 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_6aa6660a_18f1de9d6a9	April 26, 2024 9:56 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_7152ab24_18f1de91f40	April 26, 2024 9:56 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_57d62796_18f1de7d45e	April 26, 2024 9:54 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_2fe83474_18f1de358ec	April 26, 2024 9:49 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_340195ed_18f1de22944	April 26, 2024 9:48 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4802a4dd_18f1de11754	April 26, 2024 9:47 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_6d82ac0a_18f1ddfb834	April 26, 2024 9:45 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_453851fe_18f1dde242b	April 26, 2024 9:44 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_62b7c45d_18f1ddc8277	April 26, 2024 9:42 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_5f62f6bf_18f1ddb6df9	April 26, 2024 9:41 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_53252950_18f1ddb3f4c	April 26, 2024 9:41 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4a170b15_18f1dd92dce	April 26, 2024 9:38 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_21360582_18f1dd2f386	April 26, 2024 9:31 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_5b0d5180_18f1dd295fa	April 26, 2024 9:31 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_23220022_18f1dcb9e92	April 26, 2024 9:23 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_4bfdbfdb_18f1dbdb8f7	April 26, 2024 9:08 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_3b6d5d67_18f1db91d70	April 26, 2024 9:03 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_79a5dae5_18f1db87d97	April 26, 2024 9:03 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_4f86521a_18f1db807a7	April 26, 2024 9:02 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_7bf7663d_18f1db4e9af	April 26, 2024 8:59 PM	manjot.singh@sjsu.edu	QUERY	-	-	
bquxjob_33f6cb6_18f1dad3556	April 26, 2024 8:50 PM	manjot.singh@sjsu.edu	LOAD	-	-	
bquxjob_281c4455_18f1dac8601	April 26, 2024 8:49 PM	manjot.singh@sjsu.edu	LOAD	-	-	

Includes DB Connectivity / API calls

Python:

C: > Users > manjo > OneDrive > Documents > Homework > DATA 225 > MySQL connection - Updated.py > ...

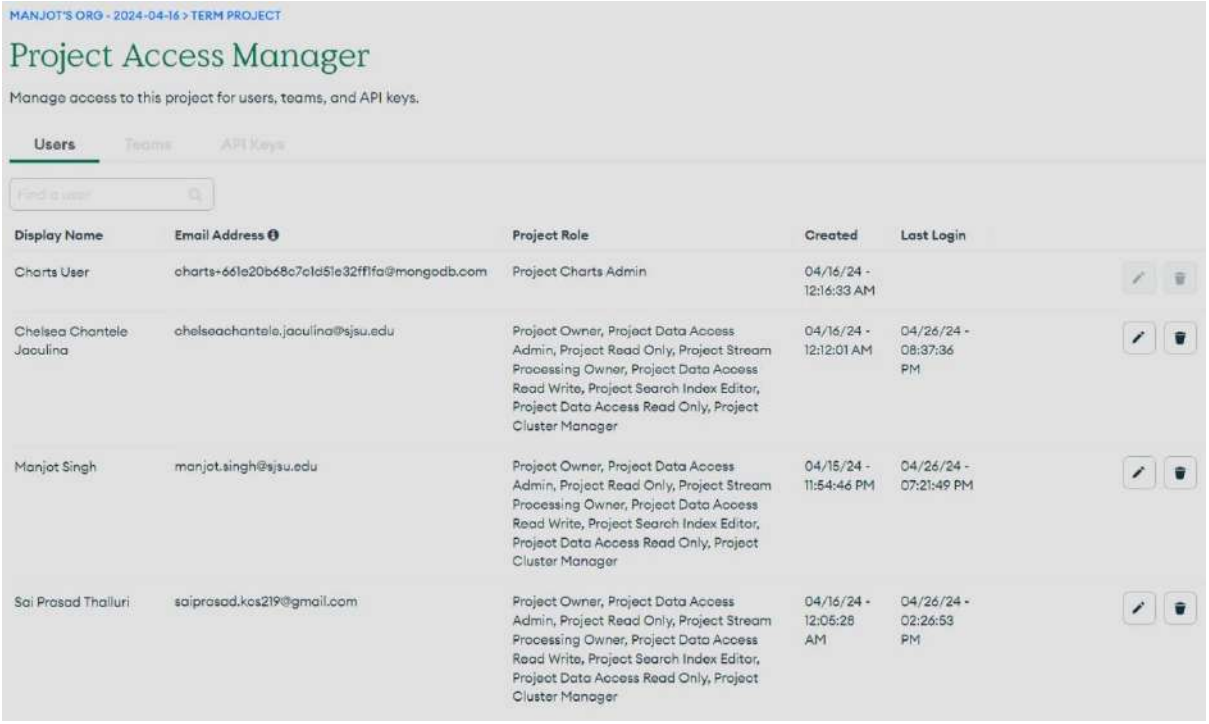
```python
import mysql.connector
import pandas as pd

# Database Credentials
host = "127.0.0.1"
user = "root"
password = "Data225DB"

# Connect to MySQL server
try:
    connection = mysql.connector.connect(host=host, user=user, password=password)
    print("Connected to MySQL server successfully!")
except mysql.connector.Error as err:
    print(f"Error connecting to MySQL server: {err}")
    exit()

cursor = connection.cursor()

# Database name
database_name = "Sleep"

# Checking if database exists
try:
    cursor.execute(f"SHOW DATABASES LIKE '{database_name}'")
    database_exists = cursor.fetchone() is not None
except mysql.connector.Error as err:
    print(f"Error checking for database: {err}")
    database_exists = False

if not database_exists:
    try:
        cursor.execute(f"CREATE DATABASE {database_name}")
        print(f"Database '{database_name}' created successfully!")
    except mysql.connector.Error as err:
        print(f"Error creating database: {err}")
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
PS C:\Users\manjo> & C:/Users/manjo/anaconda3/envs/myenv/python.exe "c:/Users/manjo/OneDrive/Documents/Homework/DATA 225/MySQL connection - Updated.py"
Connected to MySQL server successfully!
Tables created successfully!
Fact table created successfully!
Data inserted into Person table successfully!
Data inserted into Sleep table successfully!
Data inserted into PhysicalActivity table successfully!
Data inserted into Metrics table successfully!
Data inserted into Fact table successfully!
All tables (dimension and fact) created successfully!
```

Used NOSQL

MongoDB:



```
[
  {
    $group: {
      _id: {
        AgeGroup: {
          $switch: {
            branches: [
              { case: { $and: [{ $gte: ["$Age", 20] }, { $lt: ["$Age", 30] }] }, then: "20-29" },
              { case: { $and: [{ $gte: ["$Age", 30] }, { $lt: ["$Age", 40] }] }, then: "30-39" },
              { case: { $and: [{ $gte: ["$Age", 40] }, { $lt: ["$Age", 50] }] }, then: "40-49" },
              { case: { $gte: ["$Age", 50] }, then: "50+" }
            ],
            default: "Other"
          }
        },
        SleepDisorder: "$Sleep Disorder"
      },
      Count: { $sum: 1 }
    }
  },
  {
    $project: {
      _id: 0,
      AgeGroup: "$_id.AgeGroup",
      SleepDisorder: "$_id.SleepDisorder",
      Count: 1
    }
  }
]
```

PIPELINE OUTPUT OUTPUT OPTIONS

Sample of 10 documents

```
Count : 3
AgeGroup : "20-29"
SleepDisorder : "Insomnia"
```

```
Count : 9
AgeGroup : "30-39"
SleepDisorder : "Sleep Apnea"
```

```
Count : 6
AgeGroup : "30-39"
SleepDisorder : "Insomnia"
```

```
Count : 42
AgeGroup : "40-49"
SleepDisorder : "None"
```

```
Count : 12
AgeGroup : "20-29"
SleepDisorder : "None"
```

```
Count : 15
AgeGroup : "40-49"
SleepDisorder : "Sleep Apnea"
```

```
Count : 50
AgeGroup : "50+"
SleepDisorder : "Sleep Apnea"
```