Assignment #1 - Virtualization and Containerization

Name: V SAI TARUN
Id: 801421332

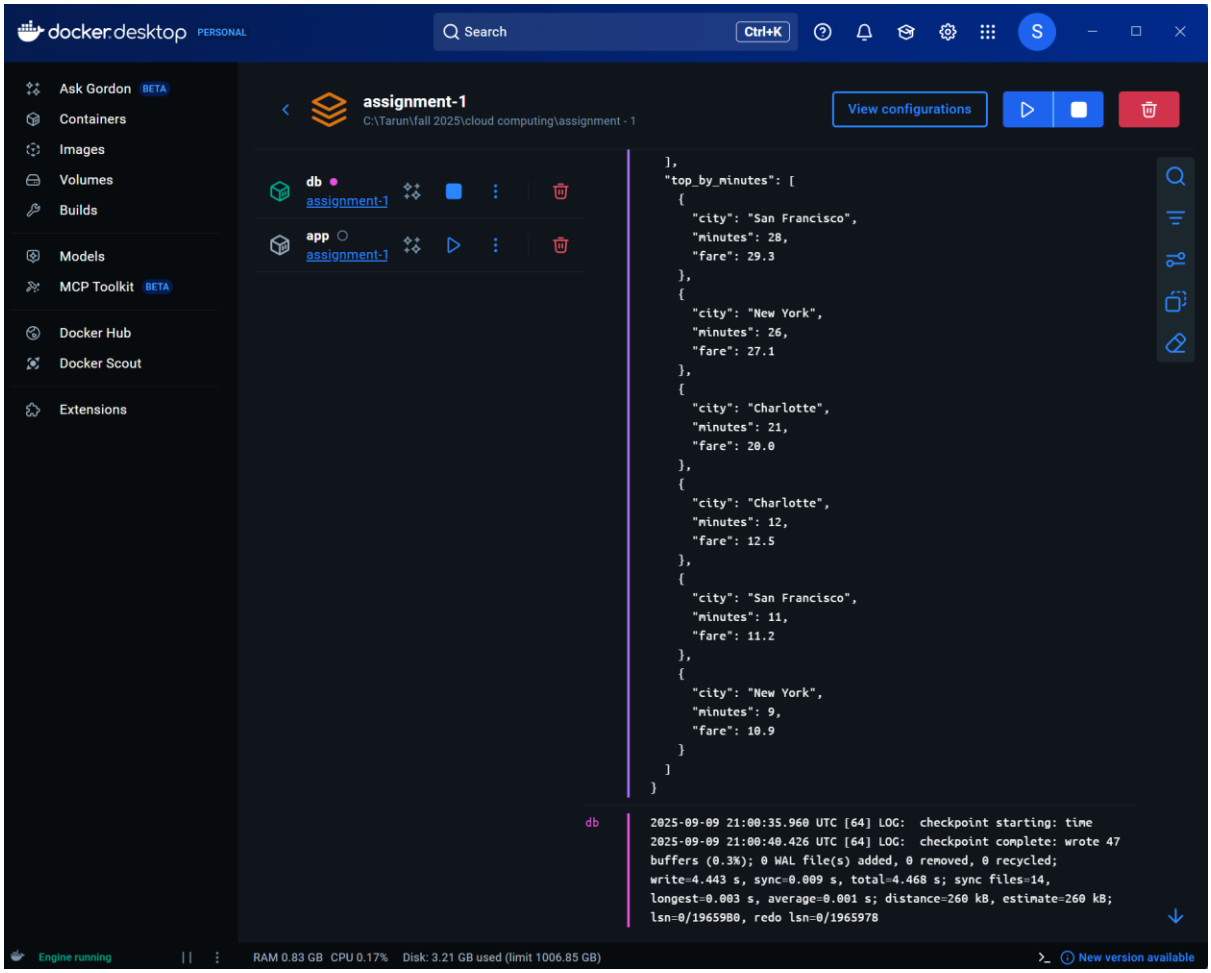
Course: Cloud Computing for Data Analysis (ITCS 6190/8190, Fall 2025)

1. Commands Used

```
# Build and start the stack
docker compose up --build

# Alternative one-command run with Makefile
make all

# Stop and remove containers + volumes
docker compose down -v
```



2. Output from Stdout (sunmmary.JSON)

```
=== Summary ===
app-1  | {
app-1  |   "total_trips": 6,
app-1  |   "avg_fare_by_city": [
app-1  |     {
app-1  |       "city": "New York",
app-1  |       "avg_fare": 19.0
app-1  |     },
app-1  |     {
app-1  |       "city": "San Francisco",
app-1  |       "avg_fare": 20.25
app-1  |     },
app-1  |     {
app-1  |       "city": "Charlotte",
app-1  |       "avg_fare": 16.25
app-1  |     }
app-1  |   ],
app-1  |   "top_by_minutes": [
app-1  |     {
app-1  |       "city": "San Francisco",
app-1  |       "minutes": 28,
app-1  |       "fare": 29.3
app-1  |     },
app-1  |     {
app-1  |       "city": "New York",
app-1  |       "minutes": 26,
app-1  |       "fare": 27.1
```

```
app-1  |    },
app-1  |    {
app-1  |      "city": "Charlotte",
app-1  |      "minutes": 21,
app-1  |      "fare": 20.0
app-1  |    },
app-1  |    {
app-1  |      "city": "Charlotte",
app-1  |      "minutes": 12,
app-1  |      "fare": 12.5
app-1  |    },
app-1  |    {
app-1  |      "city": "San Francisco",
app-1  |      "minutes": 11,
app-1  |      "fare": 11.2
app-1  |    },
app-1  |    {
app-1  |      "city": "New York",
app-1  |      "minutes": 9,
app-1  |      "fare": 10.9
app-1  |    }
app-1  |  ]
app-1  | }
```

3. This same JSON was also written to out/summary.json on the host machine.

4. Reflection

This assignment helped me understand how to use Docker Compose to run multiple services together. I learned how to configure PostgreSQL with seed data and connect a Python app using environment variables. I also practiced writing SQL queries and handling database connection retries. The most valuable takeaway was learning how to build reproducible workflows with one command. In the future, I would add more datasets and expand the Python app to compute richer statistics.