DATA JOB ANALYTICS

DASHBOARD REPORT

Project Work Flow

1. Project objective
2. Data Cleaning with Python
3. Import Data into SQL
4. Dashboard Making
5. Collecting Insights

Project Objective

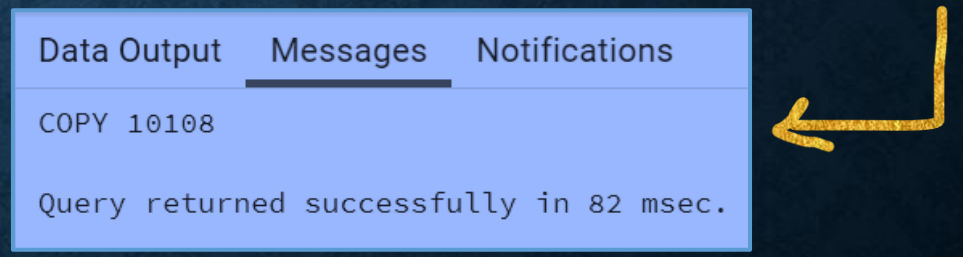
To develop a comprehensive **Job Analytics Dashboard** that provides real-time insights into key job market metrics and trends, enabling stakeholders to monitor and analyze job postings, salaries, and hiring patterns effectively.

Data Cleaning with Python

1. Converted date columns and extracted year/month for analysis
2. Removed empty rows, duplicates, and entries with missing salaries
3. Parsed job skills from strings to list format using *ast*
4. Cleaned text fields and dropped irrelevant columns

IMPORT DATA TO SQL DATABASE

1. Prepare csv file
2. Create tables in SQL
3. import csv file into SQL



NOTE: Find all SQL queries & project data- https://github.com/swapnil2468/Data-Job-Analytics-Full-Stack-Project

SQL QUERIES

```sql
CREATE DATABASE data_jobs;

-- 1. Create jobs table
CREATE TABLE jobs (
    id SERIAL PRIMARY KEY,
    job_title_short TEXT,
    job_title TEXT,
    job_location TEXT,
    job_via TEXT,
    job_schedule_type TEXT,
    job_work_from_home BOOLEAN,
    search_location TEXT,
    job_posted_date TIMESTAMP,
    job_no_degree_mention BOOLEAN,
    job_health_insurance BOOLEAN,
    job_country TEXT,
    salary_rate TEXT,
    salary_year_avg NUMERIC,
    company_name TEXT,
    job_skills TEXT,
    job_type_skills TEXT,
    year INT,
    monthname TEXT);
```

SQL QUERIES

```
-- 3. Copy csv data into SQL

COPY jobs (job_title_short, job_title, job_location, job_via,
job_schedule_type, job_work_from_home, search_location, job_posted_date,
job_no_degree_mention, job_health_insurance, job_country, salary_rate,
salary_year_avg, company_name, job_skills, job_type_skills, year, monthname)
FROM 'C:\Users\91987\Desktop\New folder (2)\Full Stack
Project\data_jobs_cleaned.csv'
DELIMITER ','
CSV HEADER;
```

TOP 5 SQL QUERIES

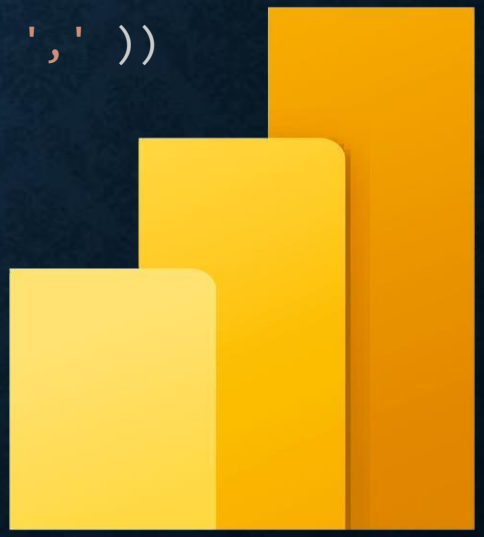
1. Most in demand skill

```sql
WITH skills_exploded AS (
    SELECT
REGEXP_REPLACE(UNNEST(string_to_array(job_skills, ',')),
'[^a-zA-Z0-9+.#-]', '', 'g') AS skill
    FROM jobs
    WHERE job_skills IS NOT NULL
    and job_title_short = 'Data Analyst'
)
SELECT LOWER(skill) AS skill, COUNT(*) AS count
FROM skills_exploded
WHERE skill IS NOT NULL AND skill <> ''
GROUP BY LOWER(skill)
ORDER BY count DESC
LIMIT 5;
```

TOP SQL QUERIES

2. Most in demand skill

```sql
WITH Skills_type AS (
    SELECT
        job_title_short,
        UNNEST(
            string_to_array(
                REGEXP_REPLACE(job_skills, '[^a-zA-Z0-9+.#,-]', '', 'g'), ',' ))
    AS skills
    FROM
        jobs)
SELECT
    count(*) FILTER(WHERE skills = 'python') as python,
    count(*) FILTER(WHERE skills = 'sql') as sql1,
    count(*) FILTER(WHERE skills = 'r') as r
FROM
    Skills_type
```



TOP SQL QUERIES

3. Top Skills by Countries

```
WITH Skills_type AS (
    SELECT job_country,
        UNNEST(string_to_array(
                REGEXP_REPLACE(job_skills, '[^a-zA-Z0-9+.#,-]','', 'g'),',' )) AS skills
    FROM
        jobs
),
Ranked AS (
    SELECT DISTINCT ON (job_country)
        job_country, skills, COUNT(*) AS skill_count
    FROM Skills_type
    GROUP BY
        job_country, skills
    ORDER BY
        job_country, skill_count DESC )
```

TOP SQL QUERIES

3. Top Skills by Countries

```
SELECT *
FROM
    Ranked
ORDER BY
    skill_count DESC
LIMIT 10;
```

TOP SQL QUERIES

4. Highest Avg. Salary of Most In demand Skills

```
WITH Skills_type AS (
    SELECT
        salary_year_avg,
        UNNEST(
            string_to_array(
                REGEXP_REPLACE(job_skills, '[^a-zA-Z0-9+.#,-]', '', 'g'), ','
            )
        ) AS skills
    FROM
        jobs
)
```

TOP SQL QUERIES

4. Highest Avg. Salary of Most In demand Skills

```sql
SELECT
    ROUND(AVG(salary_year_avg), 2) AS Average,
    TRIM(skills) AS skills,
    COUNT(skills) AS skill_count
FROM Skills_type
GROUP BY skills
HAVING COUNT(skills) > 100
ORDER BY Average DESC,skill_count DESC;
```

TOP SQL QUERIES

5. Full Time vs Part Time

```sql
WITH schedule_type AS (
    SELECT salary_year_avg,
        UNNEST(
            string_to_array(
                REGEXP_REPLACE(job_schedule_type, '[,/]',' and ', 'g'),
                ' and '
            )
        ) AS sched_type
    FROM jobs
)
SELECT
    round(AVG(salary_year_avg)
    FILTER (WHERE TRIM(sched_type) = 'Full-time'), 2) AS Full,
    round(AVG(salary_year_avg)
    FILTER (WHERE TRIM(sched_type) = 'Part-time'), 2) AS Part
FROM schedule_type;
```

DAX QUERIES

```
MostCommonJobTitle =
VAR MostFrequentJob =
    TOPN(
        1,
        SUMMARIZE(
            'public jobs',
            'public jobs'[job_title_short],
            "Count", COUNT('public jobs'[job_title_short])
        ),
        [Count],
        DESC
    )
RETURN
    CONCATENATEX(MostFrequentJob, 'public jobs'[job_title_short], ", ")
```

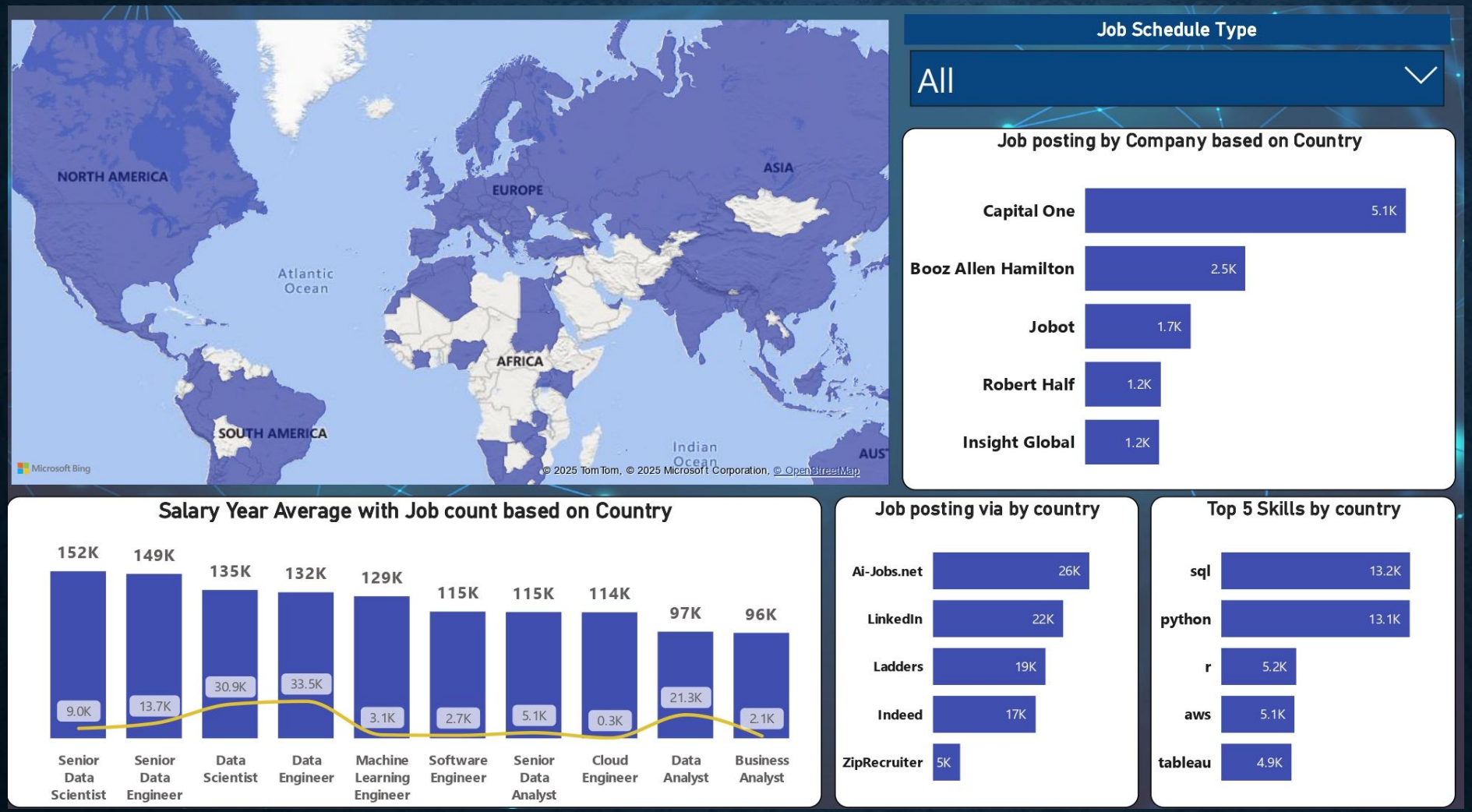
DAX QUERIES

```
Remote Jobs % =
VAR TotalJobs = COUNT('public jobs'[job_title_short])
VAR RemoteJobs = CALCULATE(COUNT('public jobs'[job_title_short]), 'public
jobs'[job_work_from_home] = TRUE)
RETURN DIVIDE(RemoteJobs, TotalJobs, 0)
```

```
monthNumber = MONTH('public jobs'[job_posted_date])
```

POWER BI DASHBOARD

POWER BI DASHBOARD



PROJECT INSIGHTS

- Total Jobs are around 121.8K
- Remote jobs make up around 15.48% of total job listings
- Overall Average Salary was around 128.1K per Anum
- Most Common Job is Data Engineer
- Job postings dropped by 50% from August to September
- Capital One posted the highest number of jobs.
- Air Job Net and LinkedIn were among the top platforms for job postings.
- SQL and Python were among the Top 5 Skills
- The US had the highest number of job postings.