

Day 9: Final Insights & Report

Dataset Used: cleaned_DataScientist.csv

Objective:

To summarize key findings from the data analysis of Data Scientist job listings across various dimensions such as job titles, skills, companies, and locations.

Key Insights:

1. Most In-Demand Skills

- Top technical skills:
 - data, machine, learning, python, sql, analytics, statistics
- Cloud & Big Data tools:
 - aws, spark, hadoop, powerbi, tableau
- Soft skills frequently mentioned:
 - communication, leadership, problem-solving

2. High-Growth Job Titles

- Top 3 titles by frequency:
 - data scientist, data analyst, data engineer
- Other emerging titles:
 - ML engineer, AI scientist, Analytics Specialist
- Synonym grouping and NLP helped normalize variations

3. Regional Job Hubs

- Top cities for data jobs:
 - New York, NY, Austin, TX, Chicago, IL, Houston, TX, Dallas, TX
- Clustering suggests Texas and East Coast are hiring hotspots

4. Top Hiring Companies

- Most active recruiters:
 - IBM, Apple, Amazon, Staffigo, Diverse Lynx
- Companies like Facebook, Humana, and CyberCoders are also significant


5. Industry-Wise Skill Patterns

- Internet & IT Services prioritize:
 - analytics, statistics, machine learning, python
- Industries like staffing & outsourcing and marketing also show demand for sql, excel, and visualization tools

6. Skill Trends by Job Level

- **Entry Level:** Emphasis on data, excel, and sql
- **Mid-Level:** Broader tools like machine learning, python, analytics
- **Senior Roles:** Add skills like leadership, strategic thinking, and team management

Final Takeaways:

- **Data roles** are evolving with hybrid requirements: tech + business
- Core skills like SQL, Python, and Data Visualization are essential.  What can I help you build?
- There's an increasing demand for **cloud**, **big data**, and **data storytelling** skills.
- **Job titles and required skills** vary significantly across **company type**, **region**, and **level**.

