

Roadmap to Discord Data Engineer Role (3rd Year B.Tech to Placement)

Prepared for: TARUN Soni\ **Target Role:** Data Engineer at Discord\ **Timeframe:** 12 Months (3rd Year to Early 4th Year)\ **Outcome:** Portfolio + Skills to Clear Discord Interviews



Month-by-Month Detailed Roadmap with Weekly Tasks



Month 1: SQL Mastery

Weeks 1-4 Tasks:

- **Week 1-2:**
 - Learn Advanced SQL (Window Functions, CTEs, Joins)
 - Practice on LeetCode / HackerRank SQL problems
 - **Week 3:**
 - Study Indexing, Performance Optimization, Query Profiling
 - **Week 4:**
 - Begin **Project 1: YouTube Analytics Data Pipeline**
 - Set up PostgreSQL, connect APIs
-



Month 2: Python for Data Engineering

Weeks 5-8 Tasks:

- **Week 5-6:**
 - Master Python libraries: pandas, pathlib, json
 - Build small ETL scripts: CSV → PostgreSQL
 - **Week 7:**
 - Linux basics: crontab, bash, permissions
 - Git basics: branches, PRs, versioning
 - **Week 8:**
 - Continue building ETL: YouTube → PostgreSQL
 - Document progress in GitHub
-

Month 3: ETL Systems & Git

Weeks 9-12 Tasks:

- **Week 9-10:**
 - Full ETL theory: Design principles, validation checks
 - Implement checks in YouTube Project (row counts, duplicates)
 - **Week 11:**
 - Use GitHub Projects for tracking progress
 - **Week 12:**
 - Finalize **YouTube Analytics Project**
 - Write README.md, draw architecture diagrams
-

Month 4: Orchestration & Transformation

Weeks 13-16 Tasks:

- **Week 13-14:**
 - Airflow basics: DAGs, scheduling, operators
 - **Week 15:**
 - DBT basics: Jinja, YAML, incremental models
 - **Week 16:**
 - **Start Project 2: Spotify Listening Data Pipeline**
 - Connect to Spotify API, PostgreSQL
-

Month 5: Cloud Data Warehousing

Weeks 17-20 Tasks:

- **Week 17-18:**
 - Learn Google Cloud, BigQuery, AWS S3 basics
 - **Week 19:**
 - Migrate Spotify project to BigQuery
 - Airflow for scheduling
 - **Week 20:**
 - Build Looker Studio dashboard
-

Month 6: Docker, Monitoring, Dashboards

Weeks 21-24 Tasks:

- **Week 21-22:**
 - Docker: containers, basics of orchestration
- **Week 23:**
 - Monitoring Airflow: logging, alerts
 - Data quality with Great Expectations

- **Week 24:**
 - Finalize **Spotify Project**
 - Documentation & screenshots for GitHub
-

Month 7: Large-Scale Data Architecture

Weeks 25-28 Tasks:

- **Week 25-26:**
 - Learn about Data Lakes, Warehouses, Schema Evolution
 - **Week 27:**
 - Kafka basics (optional, stream simulation)
 - **Week 28:**
 - **Begin Project 3: Gaming Analytics Platform**
 - Setup schemas, simulate player data
-

Month 8: Data Quality, Anomaly Detection

Weeks 29-32 Tasks:

- **Week 29-30:**
 - Deep dive: Great Expectations audits, alerts
 - **Week 31:**
 - Integrate anomaly detection into Gaming Project
 - **Week 32:**
 - Continue DBT models for KPIs (DAUs, churn)
-

Month 9: System Design for Data

Weeks 33-36 Tasks:

- **Week 33-34:**
 - Study System Design: DDIA (Ch 1-6)
 - **Week 35:**
 - Design diagrams for Gaming Project
 - **Week 36:**
 - Finalize dashboards: Looker Studio / Metabase
-

Month 10: Portfolio Polish & Communication

Weeks 37-40 Tasks:

- **Week 37-38:**
 - Write 2 technical blogs (Discord, Great Expectations)

- **Week 39:**
 - Clean GitHub, improve READMEs
 - **Week 40:**
 - Build Portfolio Site (Notion, GitHub Pages, Webflow)
-

Month 11: Interview Preparation

Weeks 41-44 Tasks:

- **Week 41-42:**
 - Mock system design interviews
 - Behavioral prep (STAR method)
 - **Week 43-44:**
 - SQL, Python LeetCode (medium difficulty)
 - Finalize Discord-fit stories
-

Month 12: Apply & Network

Weeks 45-48 Tasks:

- **Week 45-46:**
 - Apply via Discord Careers, tailor resume
 - Reach out via LinkedIn with GitHub projects
 - **Week 47-48:**
 - Review async-friendly communication prep
 - Prepare for potential interview rounds
-

(Original summary content below will be kept for final PDF, adjusted as needed...)