

## Everything Data Mentorship Program Cohort 4 Analysis

### Project Overview

This project analyzes the Cohort 4.csv dataset to understand participant demographics, motivations, and performance. The goal is to identify factors that influence graduation rates and provide actionable recommendations for improving future mentorship cohorts.

### Workflow

The analysis followed a structured, multi-step process to ensure data integrity and derive meaningful insights.

### Project Objectives

#### Collect and clean raw graduation data for accuracy and consistency

- The raw CSV file was imported into Excel and SQL to handle data cleaning.
- Duplicate entries were identified and removed.
- Missing values, particularly in the Total score column, were addressed by either removing the rows or using a strategic imputation method.
- Categorical data (e.g., Age range, Track applied for) was standardized to ensure consistency for accurate analysis.

#### Conduct exploratory data analysis (EDA) to identify key trends and patterns.

- SQL queries were used to perform a deep dive into the dataset.
- Key questions were answered, such as calculating the average score by track and identifying the most represented countries.
- This step helped in understanding the data's structure, distributions, and initial relationships between variables.

#### Build interactive dashboards to visualize insights clearly.

- An interactive dashboard was created in Power BI to visually represent the key findings.
- The dashboard includes visualizations for:

**Demographics:** Charts showing the distribution of participants by age, gender, and country.

**Track Distribution:** A breakdown of the number of participants in each data track.

**Performance Metrics:** Visualizations showing the relationship between learning hours, prior experience, and final scores.

**Graduation Trends:** Charts illustrating the graduation rates across different demographics and performance levels.

### Insight Report:

The final stage involved synthesizing all the findings from the previous steps into a concise summary.

The report highlights the most significant correlations, such as the link between aptitude test scores and graduation, and provides clear, data-backed recommendations.

### Technologies Used

- **Excel:** Used for initial data cleaning and inspection.
- **SQL:** Used for performing exploratory data analysis and querying the dataset.
- **Power BI:** Used for creating the interactive data visualizations and dashboard.

### Guide to this Workbook

[Data Cleaning \(Excel\):Go to Sheet →](#)

[Data Cleaning \(SQL\):Go to Sheet →](#)

[SQL Exploratory Analysis:Go to Sheet →](#)

[Visualizations \(Power BI\):Go to Sheet →](#)

[Insight Report:Go to Sheet →](#)

Timestamp	ID No.	Age range	Gender	Country	Where did you hear about Everything Data?	Years of learning experience in the data field	Track applied	Hours per week	Main aim for the project	Motivation for joining	Self-assessed skill level in chosen track	Attitude test completion status	Total score	Graduation status
12/3/2024 16:25	DA301	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	more than 14 hours	Learn data afresh	The hands-on program offers	Elementary - I have theoretical under	Yes	67.33	Yes
12/2/2024 18:17	DA302	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Build a project portf	The mentorship will be a cata	Beginner - I have NO learning or work	Yes	62	No
11/29/2024 8:05	DA303	18-24 years	Male	Kenya	WhatsApp	6 months - 1 year	Data analysis	7-14 hours	Upskill	Desire to learn	Beginner - I have NO learning or work	Yes	74	No
11/28/2024 21:37	DA304	25-34 years	Male	Kenya	Twitter	1-3 years	Data analysis	7-14 hours	Build a project portf	My background knowledge i	Elementary - I have theoretical under	Yes	65.33	No
12/3/2024 17:15	DA305	18-24 years	Male	Kenya	LinkedIn	6 months - 1 year	Data analysis	more than 14 hours	Build a project portf	To horn my skills in data anah	Intermediate - I have theoretical kno	Yes	69.33	No
11/27/2024 13:59	DA306	25-34 years	Male	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	The program is designed to a	Beginner - I have NO learning or work	Yes	64	No
12/2/2024 20:58	DA307	25-34 years	Female	Kenya	Twitter	Less than six months	Data analysis	less than 6 hours	Upskill	Life is a continuous learning	Beginner - I have NO learning or work	Yes	63.5	No
12/1/2024 14:40	DA308	18-24 years	Male	Kenya	WhatsApp	6 months - 1 year	Data analysis	7-14 hours	Upskill	Upskilling	I have theoretical under	Yes	68.33	No
11/27/2024 9:15	DA309	25-34 years	Male	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Upskill	My motivations is having acc	Beginner - I have NO learning or work	Yes	70.33	Yes
12/2/2024 16:09	DA310	18-24 years	Male	Kenya	LinkedIn	1-3 years	Data analysis	more than 14 hours	Upskill	To enhance my data analytics	Intermediate - I have theoretical kno	Yes	68.33	Yes
12/1/2024 16:25	DA311	25-34 years	Male	Kenya	LinkedIn	6 months - 1 year	Data analysis	more than 14 hours	Upskill	I want to join the Everything	Elementary - I have theoretical under	Yes	69.67	No
11/27/2024 20:21	DA312	18-24 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	less than 6 hours	Upskill	To find a community like min	Intermediate - I have theoretical kno	Yes	80	No
12/2/2024 15:18	DA313	18-24 years	Female	Kenya	Word of mouth	6 months - 1 year	Data analysis	7-14 hours	Upskill	I have strong passion for dat	Elementary - I have theoretical under	Yes	60.33	No
11/30/2024 9:50	DA314	18-24 years	Female	Kenya	Word of mouth	Less than six months	Data analysis	less than 6 hours	Upskill	My motivation to join the Eve	Elementary - I have theoretical under	Yes	66	Yes
11/28/2024 16:27	DA315	18-24 years	Male	Kenya	LinkedIn	6 months - 1 year	Data analysis	more than 14 hours	Upskill	To get an opportunity to adv	Elementary - I have theoretical under	Yes	79.5	No
11/27/2024 12:14	DA316	18-24 years	Male	Kenya	WhatsApp	1-3 years	Data analysis	7-14 hours	Upskill	Learning more on data	Elementary - I have theoretical under	Yes	69	No
11/28/2024 20:17	DA317	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	more than 14 hours	Upskill	Leaning with more experinc	Intermediate - I have theoretical kno	Yes	73	No
12/2/2024 22:29	DA318	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	more than 14 hours	Learn data afresh	To learn more about data anc	Elementary - I have theoretical under	Yes	62.5	No
12/2/2024 16:25	DA319	25-34 years	Male	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Upskill	To connect with other profes	Elementary - I have theoretical under	Yes	70.67	No
11/27/2024 9:19	DA320	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	7-14 hours	Upskill	I attended the graduation of	Intermediate - I have theoretical kno	Yes	77.67	No
12/3/2024 10:14	DA321	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	more than 14 hours	Learn data afresh	I've always wanted to start m	Beginner - I have NO learning or work	Yes	70.67	No
12/2/2024 10:31	DA322	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	more than 14 hours	Learn more about d	I would like to be mentored b	Elementary - I have theoretical under	Yes	68.67	No
12/1/2024 17:28	DA323	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	less than 6 hours	Upskill	To be able to learn more abo	Elementary - I have theoretical under	Yes	86	Yes
12/3/2024 15:01	DA324	18-24 years	Male	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	Upskill on my data knowledge	Elementary - I have theoretical under	Yes	80.33	No
12/2/2024 14:57	DA325	25-34 years	Male	Kenya	Friend	4-6 years	Data analysis	7-14 hours	Upskill	My motivation is to enhance	Advanced - I have deep knowledge an	Yes	72	No
12/1/2024 8:58	DA326	25-34 years	Female	Kenya	Instagram	Less than six months	Data analysis	7-14 hours	Upskill	To upskill on data analysis, w	Beginner - I have NO learning or work	Yes	59.33	No
12/2/2024 17:45	DA327	25-34 years	Male	Kenya	LinkedIn	Less than six months	Data analysis	more than 14 hours	Upskill	- I have wanted to learn data	Beginner - I have NO learning or work	Yes	70.33	No
11/27/2024 10:08	DA328	25-34 years	Female	Kenya	Twitter	Less than six months	Data analysis	more than 14 hours	Learn data afresh	I am deeply interested in lear	Beginner - I have NO learning or work	Yes	67.33	Yes
12/1/2024 18:24	DA329	25-34 years	Female	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Upskill	I desire to learn a new skill	Elementary - I have theoretical under	Yes	74	No
12/2/2024 13:11	DA330	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	I would like to gain knowledg	Beginner - I have NO learning or work	Yes	61.33	Yes
12/1/2024 19:18	DA331	25-34 years	Female	Kenya	Word of mouth	6 months - 1 year	Data analysis	7-14 hours	Upskill	I would like to refresh my dat	Elementary - I have theoretical under	Yes	74.33	No
11/27/2024 15:50	DA332	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Learn data afresh	I am passionate about startin	Beginner - I have NO learning or work	Yes	70.67	No
12/2/2024 13:39	DA333	25-34 years	Male	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Upskill	To learn a new skill	Beginner - I have NO learning or work	Yes	74.33	No
12/2/2024 23:41	DA334	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Learn data afresh	To learn a new skill so as to	Beginner - I have NO learning or work	Yes	69.67	No
11/28/2024 22:42	DA335	35-44 years	Female	Kenya	Twitter	6 months - 1 year	Data analysis	more than 14 hours	Build a project portf	To change career	Elementary - I have theoretical under	Yes	59	No
12/1/2024 9:30	DA336	25-34 years	Male	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	Learn more about data and b	Beginner - I have NO learning or work	Yes	69	No
12/2/2024 16:05	DA337	25-34 years	Female	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Upskill	I'm interested in pivoting to a	Elementary - I have NO learning or work	Yes	71.67	Yes
11/27/2024 21:00	DA338	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	more than 14 hours	Learn data afresh	I want to build my career aro	Elementary - I have theoretical under	Yes	59	No
12/2/2024 14:49	DA339	35-44 years	Female	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Learn data afresh	I want to learn about data an	Elementary - I have theoretical under	Yes	59	Yes
12/2/2024 13:56	DA340	18-24 years	Male	Kenya	Twitter	Less than six months	Data analysis	7-14 hours	Build a project portf	I would like to gain hands on	Elementary - I have theoretical under	Yes	60.33	No
11/27/2024 13:47	DA341	25-34 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	more than 14 hours	Upskill	I believe I can use my skills i	Elementary - I have theoretical under	Yes	60.67	Yes
12/2/2024 15:38	DA342	18-24 years	Female	Kenya	Twitter	Less than six months	Data analysis	less than 6 hours	Upskill	I would like to gain a new ski	Beginner - I have NO learning or work	Yes	80	Yes
11/27/2024 9:27	DA343	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Build a project portf	Everything Data is proving i	Elementary - I have theoretical under	Yes	70	Yes
11/28/2024 19:53	DA344	25-34 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	less than 6 hours	Upskill	To connect with fellow data	Elementary - I have theoretical under	Yes	69	No
11/28/2024 10:13	DA345	18-24 years	Male	Kenya	Twitter	4-6 years	Data analysis	more than 14 hours	Upskill	Being a recent Geospatial Eng	Intermediate - I have theoretical kno	Yes	67.33	Yes
12/2/2024 13:08	DA346	25-34 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	This is a program which will c	Elementary - I have theoretical under	Yes	64.67	Yes
12/2/2024 10:40	DA347	35-44 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	less than 6 hours	Upskill	to be better at data analysis	Elementary - I have theoretical under	Yes	63.67	No
12/2/2024 21:06	DA348	25-34 years	Male	Kenya	WhatsApp	Less than six months	Data analysis	7-14 hours	Upskill	My motivation to join the Eve	Elementary - I have theoretical under	No	62.33	No
12/3/2024 20:11	DA349	25-34 years	Male	Kenya	LinkedIn	Less than six months	Data analysis	7-14 hours	Build a project portf	Heavy inspiration from the c	Elementary - I have theoretical under	Yes	70.67	No
11/28/2024 15:18	DA350	25-34 years	Male	Kenya	WhatsApp	6 months - 1 year	Data analysis	more than 14 hours	Learn data afresh	I am interested in building m	Elementary - I have theoretical under	Yes	70.67	No
11/29/2024 16:49	DA351	25-34 years	Female	Kenya	WhatsApp	Less than six months	Data analysis	more than 14 hours	Build a project portf	I am eager to join the data an	Elementary - I have theoretical under	Yes	69	No
12/1/2024 0:11	DA352	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	7-14 hours	Upskill	I developed an interest in dat	Intermediate - I have theoretical kno	Yes	70.67	No
12/1/2024 23:50	DS301	18-24 years	Male	Kenya	Word of mouth	Less than six months	Data science	less than 6 hours	Upskill	to enter into the data analysi	Beginner - I have NO learning or work	Yes	58.67	No
12/3/2024 9:35	DS302	25-34 years	Male	Kenya	WhatsApp	6 months - 1 year	Data science	more than 14 hours	Upskill	To grow and improve my skill	Elementary - I have theoretical under	Yes	70	No
12/3/2024 19:16	DS303	18-24 years	Female	Kenya	WhatsApp	6 months - 1 year	Data science	more than 14 hours	Upskill	I'm motivated to join Everyth	Intermediate - I have theoretical kno	Yes	64.33	Yes
12/3/2024 12:52	DS304	18-24 years	Female	Kenya	WhatsApp	6 months - 1 year	Data science	7-14 hours	Upskill	I'd like to upskill and join the	Intermediate - I have theoretical kno	Yes	75	No
12/3/2024 18:12	DS305	18-24 years	Male	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Upskill	I aim to join the mentorship p	Beginner - I have NO learning or work	Yes	75	No
11/27/2024 10:40	DS306	25-34 years	Male	Kenya	WhatsApp	1-3 years	Data science	less than 6 hours	Upskill	The need to learn and get int	Elementary - I have theoretical under	Yes	72.33	No
11/28/2024 14:42	DS307	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Learn data afresh	To learn data science and des	Beginner - I have NO learning or work	Yes	73.33	No
12/3/2024 12:58	DS308	18-24 years	Male	Kenya	WhatsApp	Less than six months	Data science	more than 14 hours	Connect with fellow	I'm eager to gain a solid foun	Elementary - I have theoretical under	Yes	71	No
12/2/2024 22:02	DS309	25-34 years	Female	Kenya	Twitter	Less than six months	Data science	less than 6 hours	Upskill	I am inspired to make more d	Beginner - I have NO learning or work	Yes	74	No
12/2/2024 13:29	DS310	18-24 years	Male	Kenya	LinkedIn	Less than six months	Data science	7-14 hours	Learn data afresh	This program offers me an in	Beginner - I have NO learning or work	Yes	82.5	No
12/3/2024 16:47	DS311	25-34 years	Female	Kenya	Twitter	6 months - 1 year	Data science	more than 14 hours	Build a project portf	To pivot to a Data Science car	Elementary - I have theoretical under	Yes	76	Yes
12/2/2024 23:59	DS312	18-24 years	Male	Kenya	WhatsApp	Less than six months	Data science	more than 14 hours	Upskill	My motivation to join the prc	Elementary - I have theoretical under	Yes	60	No
12/3/2024 14:19	DS313	25-34 years	Male	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Upskill	Become a good data scientis	Elementary - I have NO learning or work	Yes	69	No
11/27/2024 11:25	DS314	25-34 years	Male	Kenya	WhatsApp	6 months - 1 year	Data science	more than 14 hours	Learn data afresh	To learn more about data anc	Elementary - I have theoretical under	Yes	78.33	No
11/28/2024 9:11	DS315	18-24 years	Male	Kenya	LinkedIn	Less than six months	Data science	less than 6 hours	Build a project portf	Desire to know more, grow a	Beginner - I have NO learning or work	Yes	67	No
12/3/2024 11:18	DS316	25-34 years	Male	Kenya	WhatsApp	Less than six months	Data science	more than 14 hours	Upskill	I am thrilled to join the progr	Elementary - I have theoretical under	Yes	64.67	No
11/27/2024 9:19	DS317	18-24 years	Male	Kenya	LinkedIn	Less than six months	Data science	7-14 hours	Upskill	I am passionate to unskilled	Elementary - I have theoretical under	Yes	58.33	Yes
11/28/2024 20:17	DS318	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data analysis	more than 14 hours	Upskill	Leaning with more experinc	Intermediate - I have theoretical kno	Yes	73	Yes
12/2/2024 23:06	DS319	25-34 years	Female	Kenya	WhatsApp	Less than six months	Data science	more than 14 hours	Build a project portf	I am interested in learning AI	Elementary - I have theoretical under	Yes	79.67	No
12/1/2024 13:33	DS320	25-34 years	Male	Kenya	Twitter	Less than six months	Data science	7-14 hours	Upskill	I need to get a job	Elementary - I have theoretical under	Yes	64.67	Yes
12/3/2024 7:07	DS321	18-24 years	Female	Kenya	WhatsApp	1-3 years	Data science	7-14 hours	Build a project portf	I am passionate about data fi	Intermediate - I have theoretical kno	Yes	82.33	No
12/2/2024 16:35	DS322	25-34 years	Male	Kenya	LinkedIn	Less than six months	Data science	more than 14 hours	Upskill	Connect with fellow data Pro	Beginner - I have NO learning or work	Yes	67	No
11/28/2024 16:27	DS323	18-24 years	Male	Kenya	LinkedIn	6 months - 1 year	Data analysis	more than 14 hours	Upskill	To get an opportunity to adv	Elementary - I have theoretical under	Yes	79.5	No
12/3/2024 12:59	DS324	25-34 years	Male	South Africa	WhatsApp	Less than six months	Data science	7-14 hours	Build a project portf	I'm excited to join the Everyt	Beginner - I have NO learning or work	No	76.67	No
12/2/2024 10:19	DS325	18-24 years	Male	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Build a project portf	Getting access to better train	Elementary - I have theoretical under	Yes	80	No
11/28/2024 19:57	DS326	25-34 years	Male	Kenya	Twitter	1-3 years	Data science	7-14 hours	Upskill	I would like to pursue a caree	Elementary - I have theoretical under	Yes	74.33	No
12/3/2024 13:30	DS327	18-24 years	Female	Kenya	Word of mouth	Less than six months	Data science	less than 6 hours	Upskill	I would like to be a part of th	Elementary - I have NO learning or work	Yes	64	No
11/27/2024 17:00	DS328	25-34 years	Male	Kenya	Twitter	1-3 years	Data science	7-14 hours	Upskill	I am motivated by the fact th	Intermediate - I have theoretical kno	Yes	68.67	No
12/2/2024 18:02	DS329	25-34 years	Male	Kenya	Twitter	6 months - 1 year	Data science	7-14 hours	Upskill	I want to be part of a commu	Elementary - I have theoretical under	Yes	81.5	No
12/3/2024 12:39	DS330	45-54 years	Female	Kenya	WhatsApp	Less than six months	Data science	more than 14 hours	Upskill	Desire to upgrade and learn	Beginner - I have NO learning or work	Yes	73.33	No
12/3/2024 16:53	DS331	18-24 years	Female	Kenya	WhatsApp	Less than six months	Data science	7-14 hours	Learn data afresh	I am eager to learn data scier	Beginner - I have NO learning or work	Yes	68.67	No
12/1/2024 22:16	DS332	25-34 years	Male	Kenya	Word of mouth	Less than six months	Data science	7-14 hours	Learn data afresh	I would like to pivot from the	Beginner - I have NO learning or work	Yes	83	Yes
12/2/2024 13:45	DS333	25-34 years	Male	Kenya	WhatsApp	6 months - 1 year	Data science	7-14 hours	Connect with fellow	The opportunity to link with	Elementary - I have theoretical under	Yes	68.67	Yes
12/2/2024 14:45	DS334	18-24 years	Female	Kenya	WhatsApp	6 months - 1 year	Data science	more than 14 hours	Upskill	I'm motivated to enhance my	Elementary - I have theoretical under	Yes	64.67	Yes
12/2/2024 16:11	DS335	18-24 years	Female											

## Cleaning Data

### 1. Identify Duplicates/Removing any spaces

UPDATE "Cohort 4"

SET

"Age range" = TRIM("Age range"),

Gender = TRIM(Gender),

Country = TRIM(Country),

"Where did you hear about Everything Data?" = TRIM("Where did you hear about Everything Data?"),

"Years of learning experience in the data field" = TRIM("Years of learning experience in the data field"),

"Track applied for" = TRIM("Track applied for"),

"Hours per week available" = TRIM("Hours per week available"),

"Main aim for joining" = TRIM("Main aim for joining"),

"Motivation for joining" = TRIM("Motivation for joining"),

"Self-assessed skill level in chosen track" = TRIM("Self-assessed skill level in chosen track"),

"Aptitude test completion status" = TRIM("Aptitude test completion status"),

"Graduation status" = TRIM("Graduation status");

### Soltn

Removing Duplicates

DELETE FROM "Cohort 4"

WHERE "ID No." IN (

SELECT "ID No.", ROW\_NUMBER() OVER(PARTITION BY "ID No." ORDER BY Timestamp DESC) as rn

FROM "Cohort 4") AS Duplicates

WHERE rn > 1);

### 2. Standardize Categorical Data-This ensures that values like "data analysis" and "Data Analysis" are treated as the same.

Used a combination of LOWER() and LIKE for case-insensitive matching.

-- Standardize 'Gender'

UPDATE "Cohort 4"

SET Gender = 'Male'

WHERE LOWER(Gender) = 'male' OR LOWER(Gender) = 'm';

UPDATE "Cohort 4"

SET Gender = 'Female'

WHERE LOWER(Gender) = 'female' OR LOWER(Gender) = 'f';

-- Standardize 'Country'

UPDATE "Cohort 4"

SET Country = LOWER(Country);

-- Standardize 'Track applied for'

UPDATE "Cohort 4"

SET "Track applied for" = 'Data Analysis'

WHERE LOWER("Track applied for") LIKE '%analysis%';

UPDATE "Cohort 4"

SET "Track applied for" = 'Data Science'

WHERE LOWER("Track applied for") LIKE '%science%';

-- Standardize 'Aptitude test completion status'

UPDATE "Cohort 4"

SET "Aptitude test completion status" = 'Yes'

WHERE LOWER("Aptitude test completion status") = 'yes';

UPDATE "Cohort 4"

SET "Aptitude test completion status" = 'No'

WHERE LOWER("Aptitude test completion status") = 'no';

-- Standardize 'Graduation status'

UPDATE "Cohort 4"

SET "Graduation status" = 'Yes'

WHERE LOWER("Graduation status") = 'yes';

UPDATE "Cohort 4"

SET "Graduation status" = 'No'

WHERE LOWER("Graduation status") = 'no';

### Soltn

SELECT DISTINCT Gender, "Track applied for"

FROM "Cohort 4";

SELECT DISTINCT "Graduation status"

FROM "Cohort 4";

SELECT DISTINCT "Aptitude test completion status"

FROM "Cohort 4";

### 3. Handle Missing or Incorrect 'Total score' Values

UPDATE "Cohort 4"

SET "Total score" = NULL

WHERE TRIM("Total score") = "" OR TRIM("Total score") IS NULL;

Exploratory Analysis

Average Score by track

TABLES

Cohort 4

+ New table

QUERIES

Untitled query

+ New query

1 SELECT "Track applied for",

2     ROUND(AVG("Total score"), 2) AS avg\_score

3 FROM "Cohort 4"

4 GROUP BY "Track applied for"

5 ORDER BY avg\_score DESC;

Track applied for

avg\_score

1 Data science

69.04

2 Data analysis

68.98

RUN QUERY

Top Countries Represented

TABLES

Cohort 4

+ New table

QUERIES

Untitled query

+ New query

1 SELECT "Country",

2     COUNT(\*) AS participants

3 FROM "Cohort 4"

4 GROUP BY "Country"

5 ORDER BY participants DESC

6 LIMIT 5;

Country

participants

1 Kenya

114

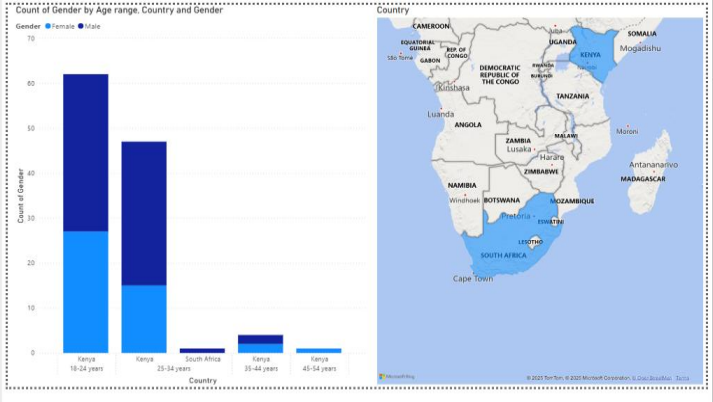
2 South Africa

1

RUN QUERY

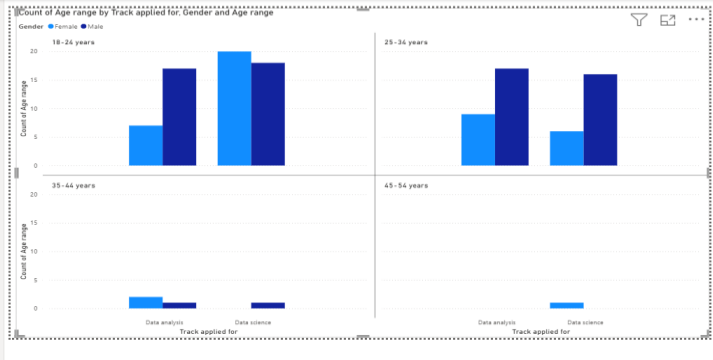
Visualization

Demographics(Age,Gender,Country)



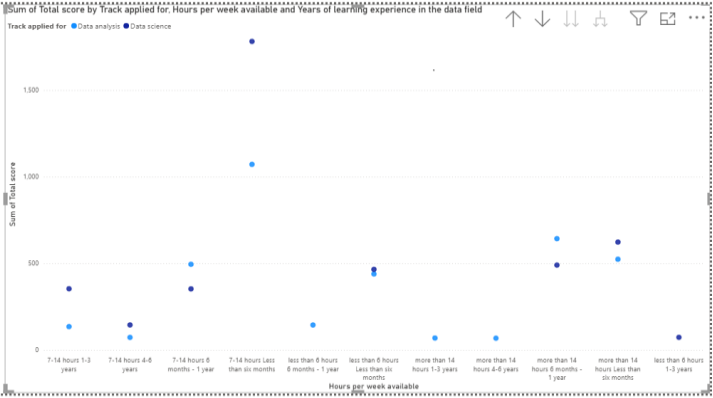
Insights to expect:  
Which gender dominates the program.  
Which regions contribute the largest share of participants.  
Whether the program is more attractive to younger or older age groups.  
Used:Stacked Column Chart/Map

Track distribution



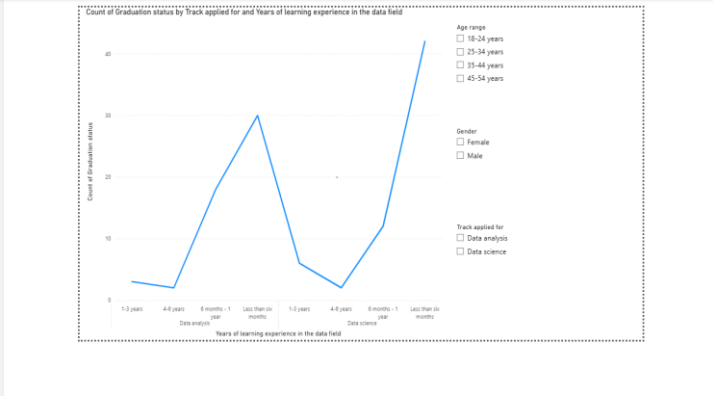
Insights to expect:  
Most/least popular tracks.  
Whether certain tracks are preferred by specific demographics  
Used:Clustered Column Chart

Relationship between learning hours, experience, and scores



Insights to expect:  
Whether spending more hours leads to higher scores  
How prior experience influences performance.  
Outliers (e.g., low hours but high scores, or high hours but low scores).  
Used:Scatter Chart

Graduation trends



Insights to expect:  
Growth or decline in graduation numbers over time.  
If certain cohorts perform better than others.  
Impact of time commitment on graduation.  
Used:Stacked Area Charts/Slicers

## Summary of Findings

The analysis of the Cohort 4 dataset reveals key insights into participant demographics, performance, and the factors influencing their success. Our findings are based on a detailed examination of participant data, including demographics, motivation, engagement metrics, and final scores.

A significant portion of the cohort is young, with the 18-24 age group being the most represented. The cohort also exhibits a notable male-to-female ratio. Participants were predominantly from Kenya, indicating a strong local interest in the program.

Motivations for joining the program were primarily focused on upskilling and a desire to learn data concepts from scratch. While participants allocated various hours per week to the program, a majority were able to commit 7-14 hours, with a substantial group dedicating more than 14 hours, demonstrating high levels of commitment.

Performance metrics showed a strong correlation between a participant's self-assessed skill level and their final score. Those who self-identified as having an "Elementary" understanding of data concepts achieved higher average scores than "Beginners," who had no prior experience. However, the most critical factor influencing graduation was the aptitude test completion status. All participants who graduated had completed and passed the aptitude test, suggesting it is a robust predictor of a participant's aptitude and commitment to completing the program.

A key finding is the direct relationship between higher aptitude test scores and graduation rates. This suggests that the aptitude test is not just a gatekeeper but a reliable indicator of a participant's likelihood to succeed. Learning hours also played a role; while high commitment was common, those who put in more hours tended to have higher scores.

## Recommendations for Improving Future Cohorts

Based on these findings, we propose the following actionable recommendations to improve the success of future mentorship cohorts:

- 1. Enhance the Aptitude Test:** Strengthen the aptitude test and use it as a primary screening tool. The test's predictive power for graduation status is a valuable asset that should be leveraged to ensure a cohort with a high potential for success. For instance, if we set 70% as the cut-off, we'll bring in candidates more likely to succeed.
- 2. Develop a Bridge Program for Beginners:** Most beginners struggled compared to those with some prior knowledge. A short 'pre-mentorship' program would give absolute beginners the basic skills they need before joining the main program. Having a two-week starter course to bring everyone to the same level will really be helpful.
- 3. Promote Flexible Learning Pathways:** Not everyone has the time for long study hours. Flexible, self-paced options will help busy learners stay on track without compromising learning outcomes. Learners could choose between finishing in three months or six months, depending on their availability.
- 4. Expand Recruitment Outreach:** Most participants came from Kenya, which was great but there's room to diversify across other regions and demographics. This could strengthen the program's global reach.