



Data Insights From Aadhaar A Comprehensive Analysis Using Qlik

Category: **Data Analytics**

Skills Required:

- **Data Analysis Skills:**
 - Understanding of data cleaning and manipulation techniques.
 - Experience with statistical analysis concept

Overview:

Description of project

This project aims to conduct a comprehensive analysis of Aadhaar data using Qlik Sense. The analysis will focus on extracting valuable insights to improve decision-making, policy formulation, and operational efficiency within the National Identity Authority of India.

Key aspects:

- Analyze demographics, enrollment trends, and activation rates of Aadhaar holders.
- Identify regional variations and potential areas for improvement.
- Utilize Qlik's visualizations to present findings in an interactive and clear format.
- Generate recommendations based on the analysis to enhance program effectiveness.

This project is ideal for showcasing your data analysis expertise and Qlik Sense skills in a real-world context.

Purpose:

- **Improve decision-making:** By analyzing Aadhaar data, the project can reveal trends and patterns that inform better decisions for the National Identity Authority of India (UIDAI). This could involve resource allocation, outreach program strategies, and policy formulation.
- **Enhance operational efficiency:** Insights from the analysis can help identify areas for improvement within the Aadhaar program. This might involve streamlining enrollment processes, reducing rejections, and optimizing resource utilization.
- **Gain deeper understanding of Aadhaar users:** Analyzing demographics, regional enrollment patterns, and activation rates provides a clearer picture of who is using Aadhaar and how. This knowledge can be used to tailor future initiatives for better program reach and impact.

Achievements:

- **Identify underserved regions:** The project can pinpoint areas with low Aadhaar penetration, allowing UIDAI to focus outreach efforts and bridge the gap.
- **Analyze enrollment rejections:** By understanding reasons for enrollment rejections (if available in the data), the project can help identify and address potential issues in the enrollment process.
- **Improve user experience:** Insights can guide improvements to the Aadhaar program, making it more efficient and user-friendly for citizens.

Data-driven policy formulation: The project's findings can be used to create data-backed policies that effectively address the evolving needs of Aadhar program

Interactive dashboards for stakeholders: Qlik Sense allows for creating interactive dashboards that empower stakeholders within UIDAI to explore the data and gain insights on their own.

Technical Architecture :

Data Source:

- The project will rely on Aadhaar data, obtained following proper data privacy regulations.
- The specific format of the data source (e.g., CSV, database) will depend on how it's provided by UIDAI or a trusted source.

Data Layer:

- **Data Acquisition:**
 - Develop scripts or tools to extract the Aadhaar data from the source.
 - Ensure data security and anonymization if necessary.
- **Data Transformation and Cleaning:**
 - Use a data warehousing or data lake solution (optional) to store and manage the data.
 - Perform data cleaning tasks like handling missing values, formatting inconsistencies, and creating derived fields using tools like Python or dedicated data wrangling software.

Analytics Layer:

- **Qlik Sense Desktop or Server:**
 - Load the cleaned Aadhaar data into Qlik Sense for analysis.
 - Utilize Qlik's data modeling capabilities to establish relationships between different data points

PROBLEM STATEMENT

SPECIFY THE BUSINESS PROBLEM: Aadhaar Card is a 12-digit unique identification number issued by the Unique Identification Authority of India (UIDAI), a statutory authority established by the Indian government. The Aadhaar initiative aims to provide a universal and robust identity infrastructure for residents of India. Analyze Aadhaar data (demographics, enrollment, location) with Qlik Sense to extract insights for UIDAI. Improve decision-making, policy, and efficiency for the Aadhaar program.

BUSINESS REQUIREMENT:

Data Access and Security:

- **Data Source:** Define the specific format and source of the Aadhaar data (e.g., database, CSV files).
- **Data Security:** Ensure strict adherence to data privacy regulations and security best practices when handling Aadhaar data. This might involve anonymization techniques and access control measures.

Data Analysis and Insights:

- **Focus Areas:** Specify the key areas of interest for analysis, such as:
 - Demographics of Aadhaar holders (age, gender, region)
 - Enrollment trends across states and districts
 - Aadhaar activation rates
 - Reasons for enrollment rejections (if available)

LITERATURE SURVEY:

A literature survey for the Aadhaar analysis would involve researching and reviewing previous studies, articles, and reports on the topic. This could include information on the methods and techniques used for analyzing Aadhaar Analysis, as well as the results and conclusions of these studies. It is recommended to explore academic databases such as PubMed, IEEE Xplore, Google Scholar, and institutional repositories. Additionally, government reports and publications can provide insights into the latest developments in Aadhaar analysis.

Government Sources:

- **UIDAI Annual Reports:** These reports provide valuable insights into Aadhaar program performance, demographics, and key initiatives. Analyze recent reports to understand current trends and challenges.
- **formats UIDAI Technical Specifications:** These documents detail the data structure and

used in the Aadhaar database. This information is crucial for data acquisition and cleaning.

News Articles and Industry Reports:

Look for recent news articles or reports covering Aadhaar enrollment trends, regional disparities, or challenges faced by the program. These can help identify areas of particular interest for your analysis.

- Explore reports from consulting firms or research institutions focusing on digital identity programs or citizen data analysis.

SOCIAL OR BUSINESS IMPACT:

Improved Decision-Making:

- **Data-driven insights:** Qlik Sense will enable UIDAI to analyze vast amounts of Aadhaar data and extract actionable insights. This data can inform strategic decisions regarding:
 - **Resource allocation:** Optimizing resource allocation based on enrollment trends, regional disparities, and program usage patterns.
 - **Policy formulation:** Data-backed evidence can guide policy development to address specific challenges and improve program effectiveness.
 - **Outreach strategies:** Insights can help identify areas requiring targeted outreach efforts to achieve wider Aadhaar adoption.

DATA COLLECTION

Downloading the Dataset with the link:

Before DataPreparation

https://docs.google.com/spreadsheets/d/1tdLMFay_8w5oelsbg72b6p8k7YzLZnbtUSZCdRrHsWU/edit?usp=drive_link

Connect Data with Qlik Sense:

<https://wjxauf17i2gonbs.sg.qlikcloud.com/>

Data Preparation:

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned we can move to visualization.

Data Cleaning and Transformation:

- **Data Cleaning:**
 - Identify and handle missing values: Decide on a strategy for missing values (e.g., imputation, exclusion).
 - Identify and correct inconsistencies: Look for typos, outliers, and formatting errors. Standardize date formats and address abbreviations.
- **Data Transformation:**
 - Derive new fields: Create calculated fields based on existing data (e.g., age from birth date).
 - Data categorization: Group data points into categories (e.g., age groups, income brackets).

link After DataPreparation:

<https://wjxauf17i2gonbs.sg.glikcloud.com/sense/app/8c69a566-9162-4907-84d9-699a1fe4b44c/datamanager/datamanager>

Qlik ... Prepare Data manager Analyze Sheet Narrate Storytelling Aadhar_dataset.csv

+ Add data Concatenate or join Associations Load data

Aadhar_dataset Columns: 15 Aadhar_dataset.csv Rows: 409275

Unpivot Add field Select data from source

Registrar	Enrolment Agency	State	District	Sub District	Pin Code	Gender	Age
Allahabad Bank	A-Onerealtors Pvt Ltd	Uttar Pradesh	Allahabad	Meja	212303	F	
Allahabad Bank	Asha Security Guard Services	Uttar Pradesh	Sonbhadra	Robertsganj	231213	M	
Allahabad Bank	SGS INDIA PVT LTD	Uttar Pradesh	Sultanpur	Sultanpur	227812	F	
Allahabad Bank	Sri Ramraja Sarkar Lok Kalyan Trust	Uttar Pradesh	Shamli	Shamli	247775	M	
Allahabad Bank	Transmoovers India	Uttar Pradesh	Gorakhpur	Sahjanwa	273001	M	
Allahabad Bank	Transmoovers India	Uttar Pradesh	Varanasi	Pindra	221101	M	
Allahabad Bank	Transmoovers India	Uttar Pradesh	Varanasi	Varanasi	221001	M	
Allahabad Bank	Transmoovers India	Uttar Pradesh	Varanasi	Varanasi	221002	M	
Allahabad Bank	Transmoovers India	Uttar Pradesh	Varanasi	Varanasi	221002	M	
Allahabad Bank	Vedavaag Systems Limited	Uttar Pradesh	Bara Banki	Nawabganj	225301	M	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Assam	Marigaon	Bhuragaon	782121	M	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Bihar	Gopalganj	Vijayeeupur	841508	M	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Karnataka	Bagalkot	Badami	587114	M	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Karnataka	Bagalkot	Badami	587155	F	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Karnataka	Bagalkot	Badami	587155	M	
Atalji Janasnehi Directorate Government of Karnataka	Atalji Janasnehi Directorate GOK	Karnataka	Bagalkot	Badami	587201	F	

Show data preview

Data Transformation:

Calculated field:

Qlik ... Prepare Data manager Analyze Sheet Narrate Storytelling Aadhar_dataset.csv

+ Add data Concatenate or join Associations Load data

Aadhar_dataset Columns: 15 Aadhar_dataset.csv Rows: 409275

Unpivot Add field Select data from source

District	Sub District	Pin Code	Gender	Age	generated	Enrolment Re...	email	mobile number	Age Gr
Allahabad	Meja	212303	F	7	1	0	0	0	1 Minor
Sonbhadra	Robertsganj	231213	M	8	1	0	0	0	0 Minor
Sultanpur	Sultanpur	227812	F	13	1	0	0	0	1 Minor
Shamli	Shamli	247775	M	6	1	0	0	0	1 Minor
Gorakhpur	Sahjanwa	273001	M	8	1	0	0	0	1 Minor
Varanasi	Pindra	221101	M	14	1	0	0	0	1 Minor
Varanasi	Varanasi	221001	M	9	1	0	0	0	1 Minor
Varanasi	Varanasi	221002	M	4	1	0	0	0	1 Minor
Varanasi	Varanasi	221002	M	19	1	0	0	0	1 Major
Bara Banki	Nawabganj	225301	M	8	1	0	0	0	0 Minor
Marigaon	Bhuragaon	782121	M	22	1	0	0	0	1 Major
Gopalganj	Vijayeeupur	841508	M	26	1	0	0	0	1 Major
Bagalkot	Badami	587114	M	27	1	0	0	0	1 Major
Bagalkot	Badami	587155	F	2	1	0	0	0	1 Minor
Bagalkot	Badami	587155	M	67	1	0	0	0	1 senior
Bagalkot	Badami	587201	F	32	1	0	0	0	1 Major
Bagalkot	Badami	587203	M	27	1	0	0	0	1 Major
Bagalkot	Badami	587206	F	40	1	0	0	0	0 Major
Bagalkot	Badami	587206	M	28	1	0	0	0	1 Major
Bagalkot	Badami	587206	M	44	1	0	0	0	1 Major

Update calculated field

Name Regions

Expression

if (Match(State, 'Maharashtra', 'Gujarat', 'Goa', 'Rajasthan', 'Dadra and Nagar Haveli', 'Daman and Diu', 'Daman', 'Diu', 'Others'), 'western',

Preview

fx

Central

Eastern

North-Eastern

Northern

Update

Show data preview

Selecting Data from Source

Aadhar_dataset.csv | Edit selections - Aadhar_dataset

Field names: Embedded field names ▼
Header size: - 0 +
Delimiter: Comma ▼
Quoting: MSQ ▼
Comment:
Character set: 28591 (ISO 8859-1 Latin I) ▼
Ignore End-Of-File character? ☐

☒ Select all fields

Filter fields

<input checked="" type="checkbox"/>	Registrar	<input checked="" type="checkbox"/>	Enrolment Agency	<input checked="" type="checkbox"/>	State	<input checked="" type="checkbox"/>	Dis...	<input checked="" type="checkbox"/>	Sub Dis...	<input checked="" type="checkbox"/>	Pin C...	<input checked="" type="checkbox"/>	Ge...	<input checked="" type="checkbox"/>	Age	<input checked="" type="checkbox"/>	gener...	<input checked="" type="checkbox"/>	Enrolment Reje...	<input checked="" type="checkbox"/>	e...	
	Allahabad Bank		A-Onerealtors Pvt Ltd		Uttar Pradesh		Allahabad		Meja		212303		F		7		1		0		0	1
	Allahabad Bank		Asha Security Guard Services		Uttar Pradesh		Sonbhadra		Robertsganj		231213		M		8		1		0		0	0
	Allahabad Bank		SGS INDIA PVT LTD		Uttar Pradesh		Sultanpur		Sultanpur		227812		F		13		1		0		0	1
	Allahabad Bank		Sri Ramraja Sarkar Lok Kalyan Trust		Uttar Pradesh		Shamli		Shamli		247775		M		6		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Gorakhpur		Sahjanwa		273001		M		8		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Pindra		221101		M		14		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Varanasi		221001		M		9		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Varanasi		221002		M		4		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Varanasi		221002		M		10		0		1		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Varanasi		221002		M		19		1		0		0	1
	Allahabad Bank		Transmoovers India		Uttar Pradesh		Varanasi		Varanasi		225301		M		8		1		0		0	0
	Allahabad Bank		Vedavaag Systems Limited		Uttar Pradesh		Bara Banki		Nawabganj		782121		M		22		1		0		0	1
	Atalji Janasnehi Directorate Govern		Atalji Janasnehi Directorate GOK		Assam		Marigaon		Bhuragaon		841508		M		26		1		0		0	1
	Atalji Janasnehi Directorate Govern		Atalji Janasnehi Directorate GOK		Bihar		Gonatalani		Mfauzenpur										0		0	1

Cancel Update

Loading the Data :

Qlik ... Prepare Data manager Analyze Sheet Narrate Storytelling Aadhar_dataset.csv

+ Add data Concatenate or join

Aadhar_dataset Columns: 15 Rows: 409275

District Sub District

Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hadagali
Ballari	Hagaribommanahalli
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Hospet
Ballari	Kudliel

mobile number Age Group Regions Aadhar Activa...

0	0	Major	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	0	senior	Southern	yes
0	1	senior	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	senior	Southern	yes
0	1	Major	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	Major	Southern	yes
0	1	senior	Southern	yes
0	1	Major	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes
0	1	Minor	Southern	yes

Your data is loading

Elapsed time: 00:00:01

Abort

Unpivot Add field Select data from source Load data

Show data preview

data preparation link:

https://drive.google.com/file/d/1zlukVR1gZ897jc-ZV_3o6PF8PbdAPPOv/view?usp=drive_link

DATA VISUALIZATION:

Data visualization is a powerful tool for communicating insights from data analysis. Here are some effective visualization techniques you can leverage in Qlik for your Aadhaar project:

Overall Enrollment Distribution:

- **Bar chart:** Show the distribution of Aadhaar enrollments across various Indian states or districts.
- **Heatmap:** Visually represent enrollment density across a geographic map of India.

Demographic Analysis:

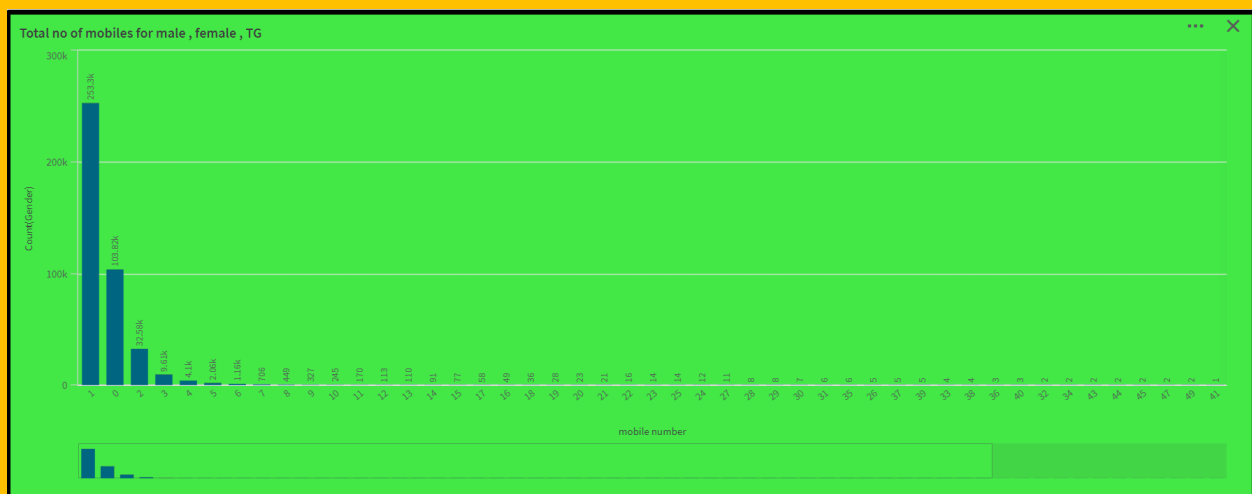
- **Histogram:** Analyze the age distribution of Aadhaar enrollees.
- **Scatter plot:** Explore correlations between age and other demographic factors like gender or location.

Aadhaar Usage Patterns:

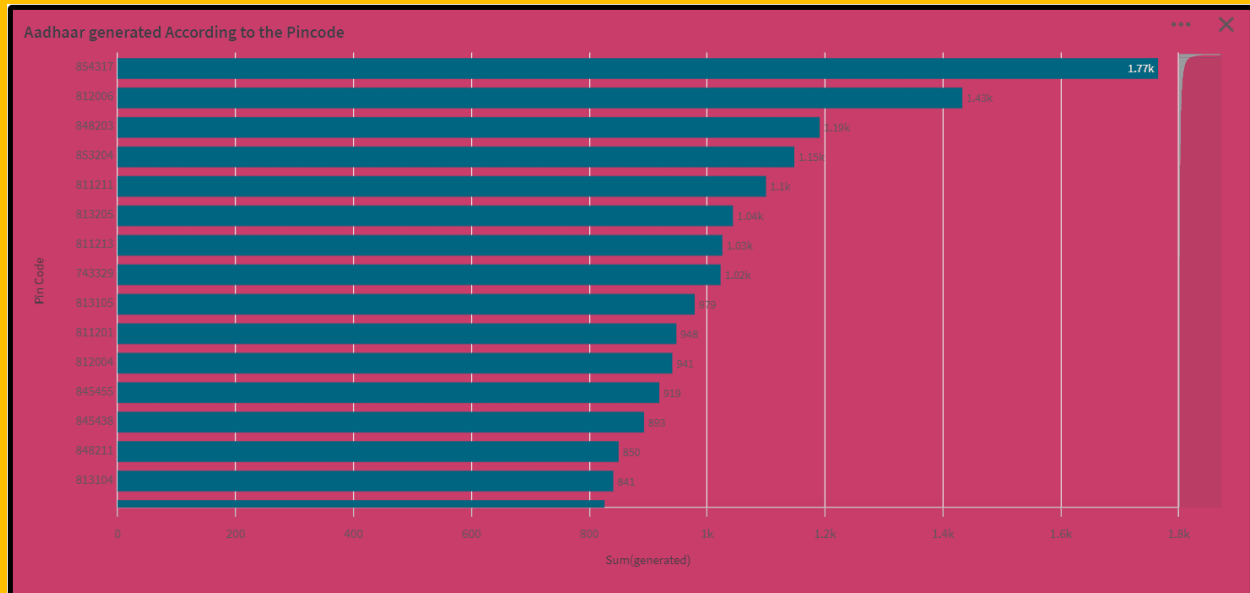
- **Line chart:** Track trends in Aadhaar authentication requests over time.
- **Pie chart:** Represent the proportion of Aadhaar usage across different sectors (e.g., banking, government services)

UNIQUE VISUALIZATIONS:

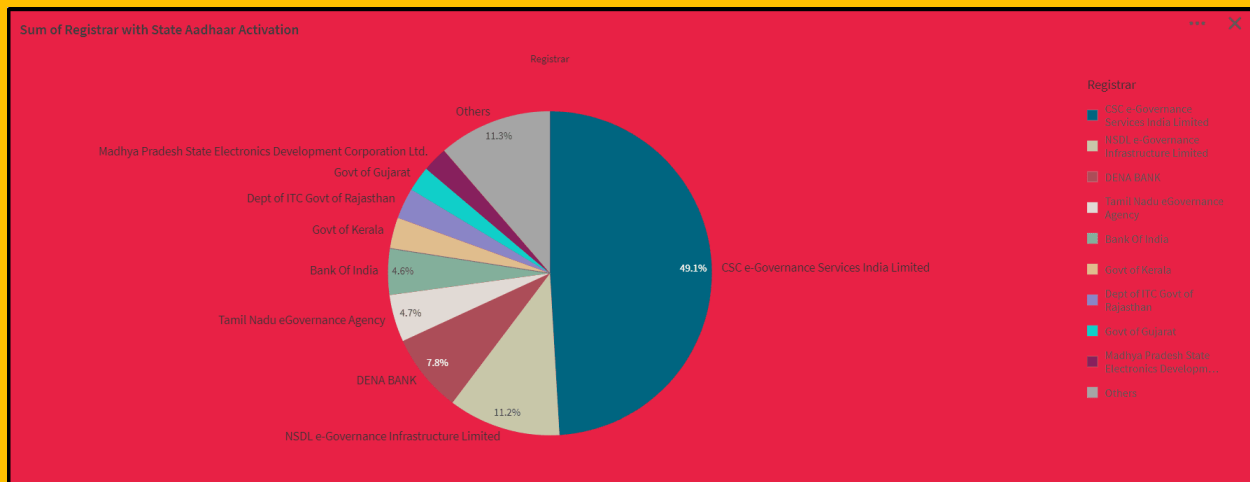
Total no of mobiles for male, female,Tg



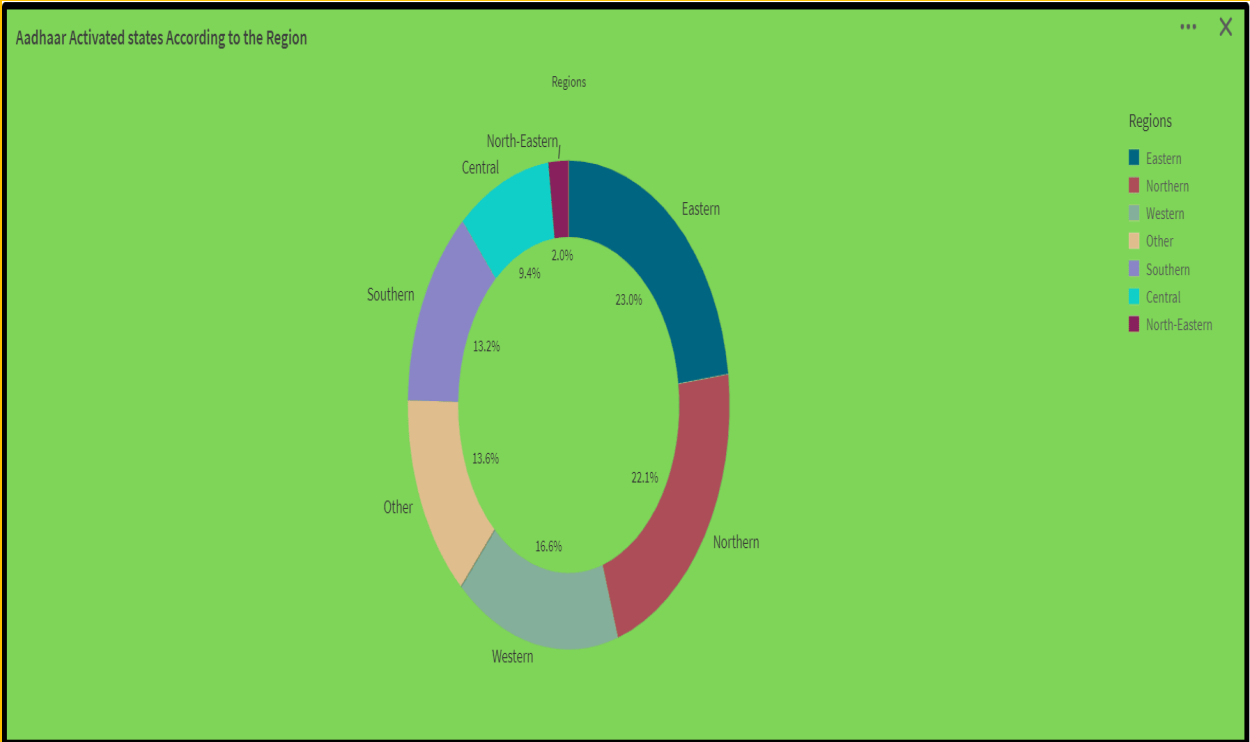
Aadhaar Generated According to the Pincode



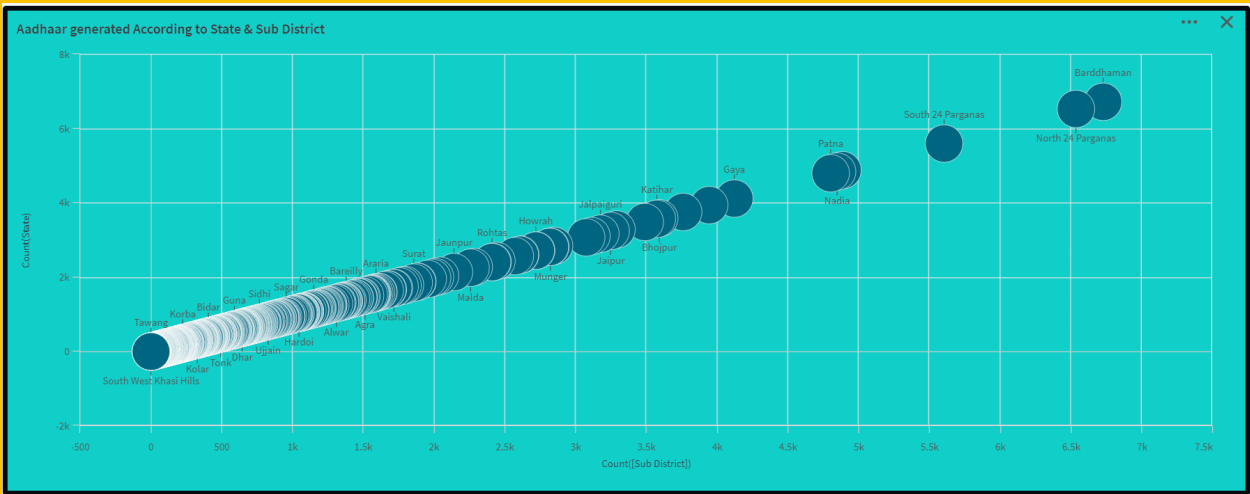
Sum of Registrar with State Aadhaar Activation



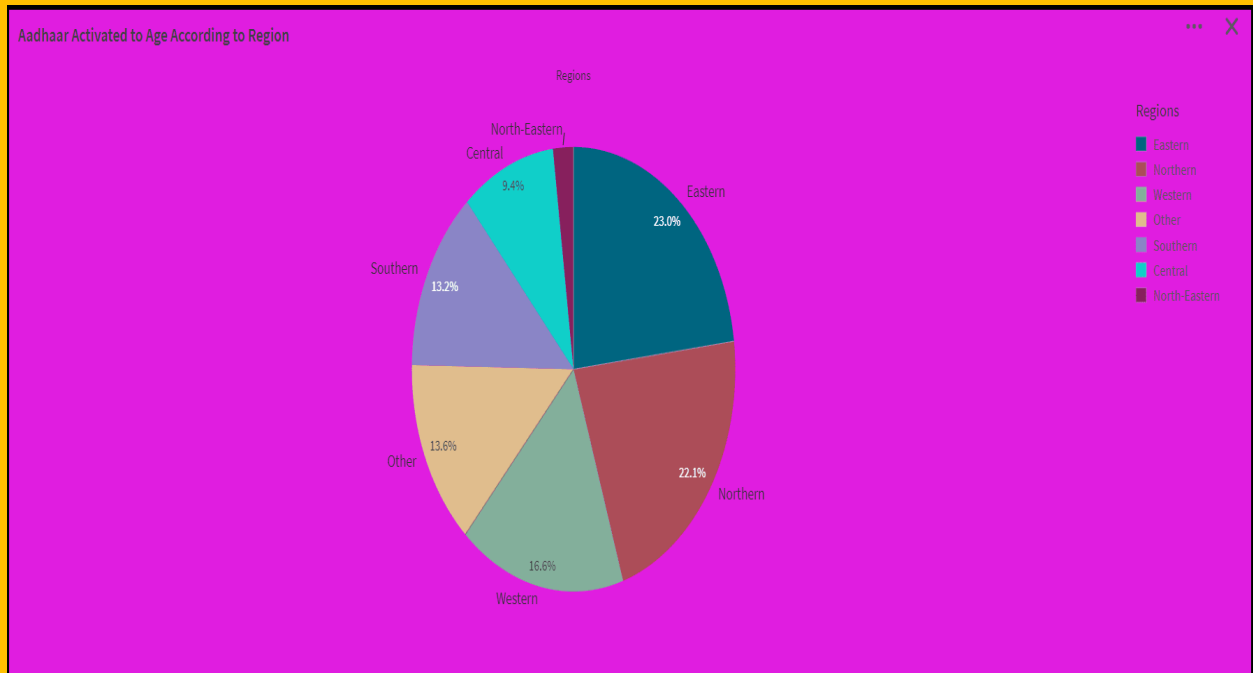
Aadhaar Activated states According to Region



Aadhaar generated According to State & sub District

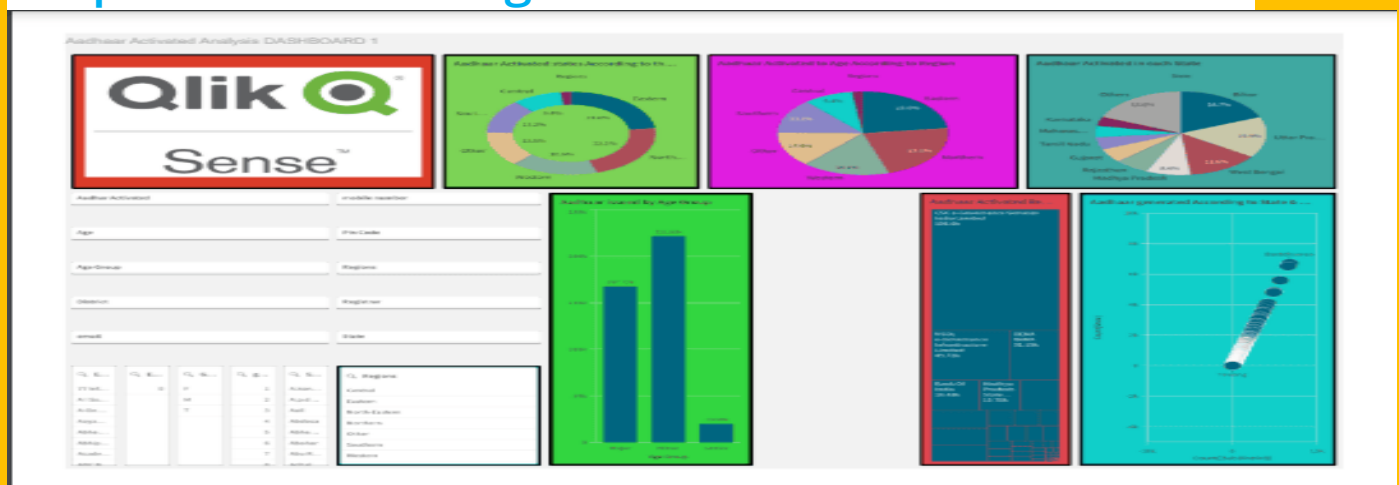


Aadhaar Activated to Age According Region

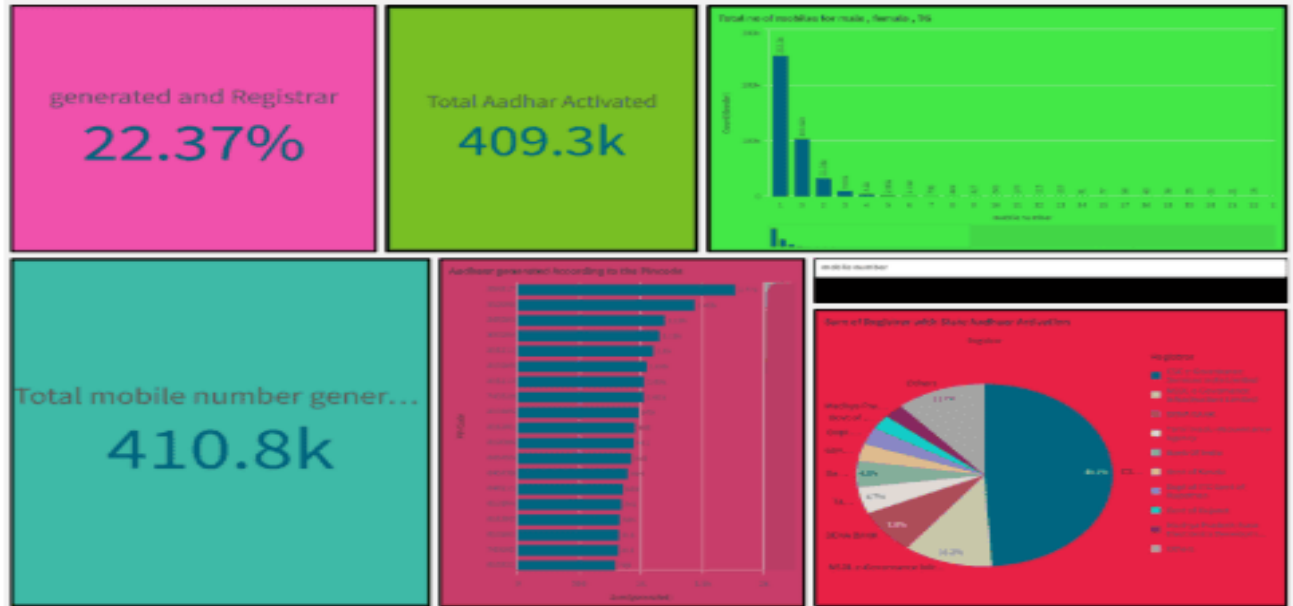


DASHBOARD

Responsive And Design Of Dashboard

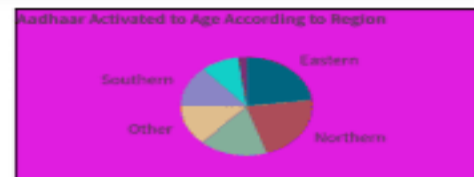


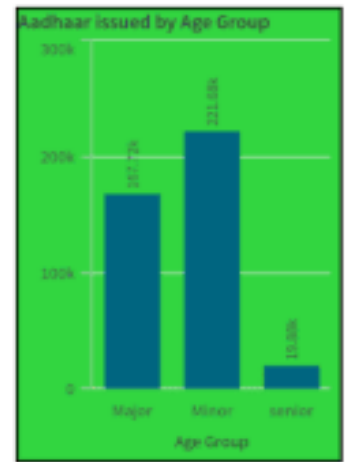
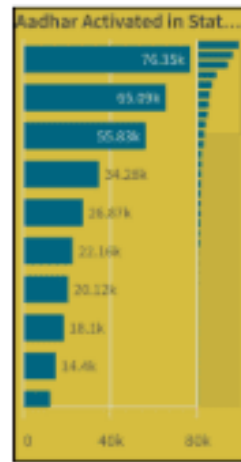
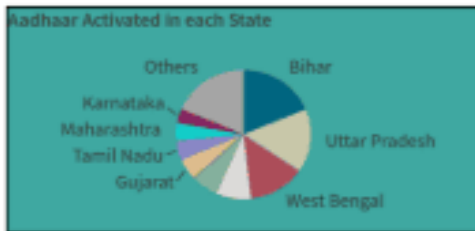
DASHBOARD 2



STORY

Design Of Story



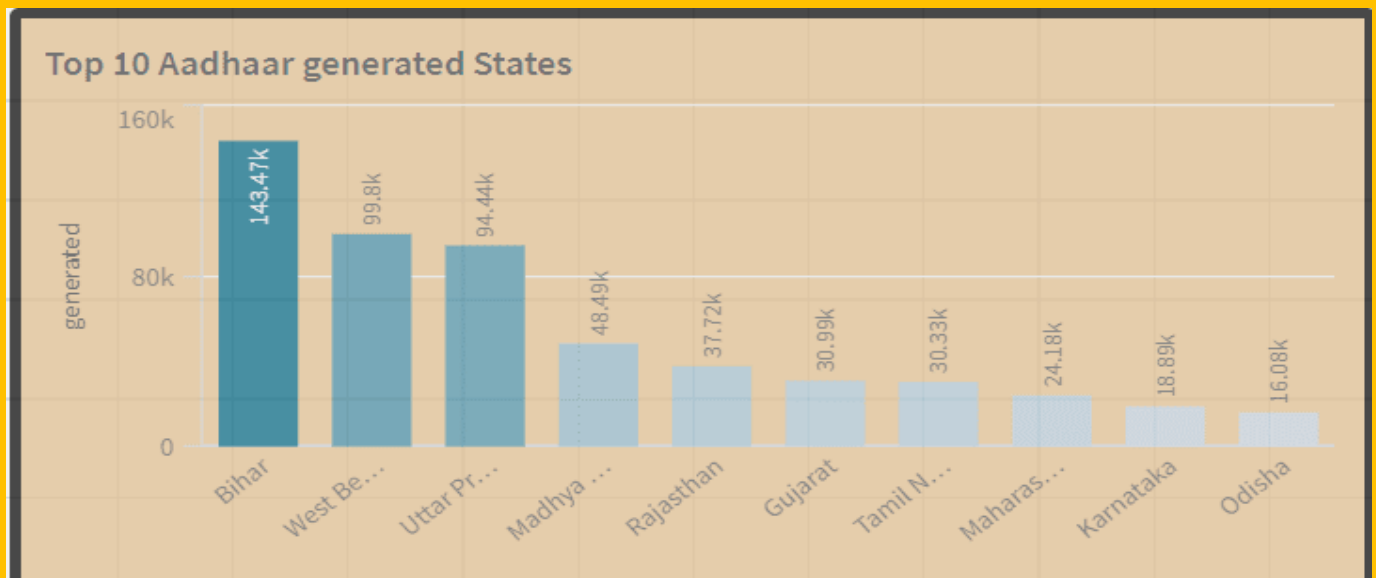


Performance Testing

Amount of Data Loaded:

Registrar, Enrollment Agency, State, District, Sub-District, Pin Code, Age, generated, email, Mobile, Age Group, Aadhaar activated, Region.

Utilization of Data filter



NO OF CALCULATION FIELD

Update calculated field

?

↶ ↷

×

Name

Age Group

Expression

If (Age > 60, 'senior',
If (Age >= 18, 'Major',
If (Age < 18, 'Minor')))

f_x

▼

▼

▼

Preview

🔍

Major

Minor

senior

Update

If (Age > 60, 'senior',

If (Age >= 18, 'Major',

If (Age < 18, 'Minor'))).

Update calculated field

?

↶ ↷

✕

Name

Regions

Expression

```
if (Match(State,
'Maharashtra', 'Gujarat',
'Goa', 'Rajasthan', 'Dadra and
Nagar Haveli', 'Daman and
Diu', 'Daman', 'Diu',
'Others'), 'Western',
```

Preview

fx

≡

- +

× ÷

🔍

Central

Eastern

North-Eastern

Northern

Update

if (Match(State, 'Maharashtra', 'Gujarat', 'Goa', 'Rajasthan', 'Dadra and Nagar Haveli', 'Daman and Diu', 'Daman', 'Diu', 'Others'), 'Western',

if (Match(State, 'Chhattisgarh', 'Madhya Pradesh'), 'Central',

if (Match (State, ' West Bengal', 'Odisha', 'Jharkhand', 'Bihar'
), 'Eastern',

if (Match (State, 'Karnataka', 'Andhra Pradesh', 'Tamil
Nadu', 'Telangana', 'Kerala', 'Andaman and Nicobar Islands',
'Puducherry', 'Lakshadweep'), 'Southern',

if (Match(State, 'Jammu and Kashmir', 'Punjab', 'Himachal
Pradesh', 'Uttarakhand', 'Haryana', 'Uttar Pradesh', 'Delhi',
'Chandigarh'), 'Northern',

if (Match(State, 'Assam', 'Meghalaya', 'Tripura', 'Nagaland',
'Arunachal Pradesh', 'Sikkim', 'Mizoram', 'Manipur'), 'North-
Eastern', 'Other'))))))).

Update calculated field

?

↗

✕

Name

Aadhar Activated

Expression

If (generated > 0, 'yes', 'No')

f_x

▼

▼

+

+

+

+

▼

Preview

🔍

yes

Update

If (generated > 0, 'yes','No')

NO OF VISUALIZATIONS /GRAPHS

- 1.Total no of mobiles for male, female,Tg**
- 2.Aadhaar Generated According to the Pincode**
- 3.Sum of Registrar with State Aadhaar Activation**
- 4.Aadhaar Activated states According to Region**
- 5.Aadhaar generated According to State & sub District**
- 6.Aadhaar Activated to Age According Region**

PROJECT DEMONSTRATION LINK:

https://drive.google.com/file/d/18c2i59366k8Hmfvjg75RopB_lg1U6zmu/view?usp=drive_link

**THANK
YOU**

DesComments.com

