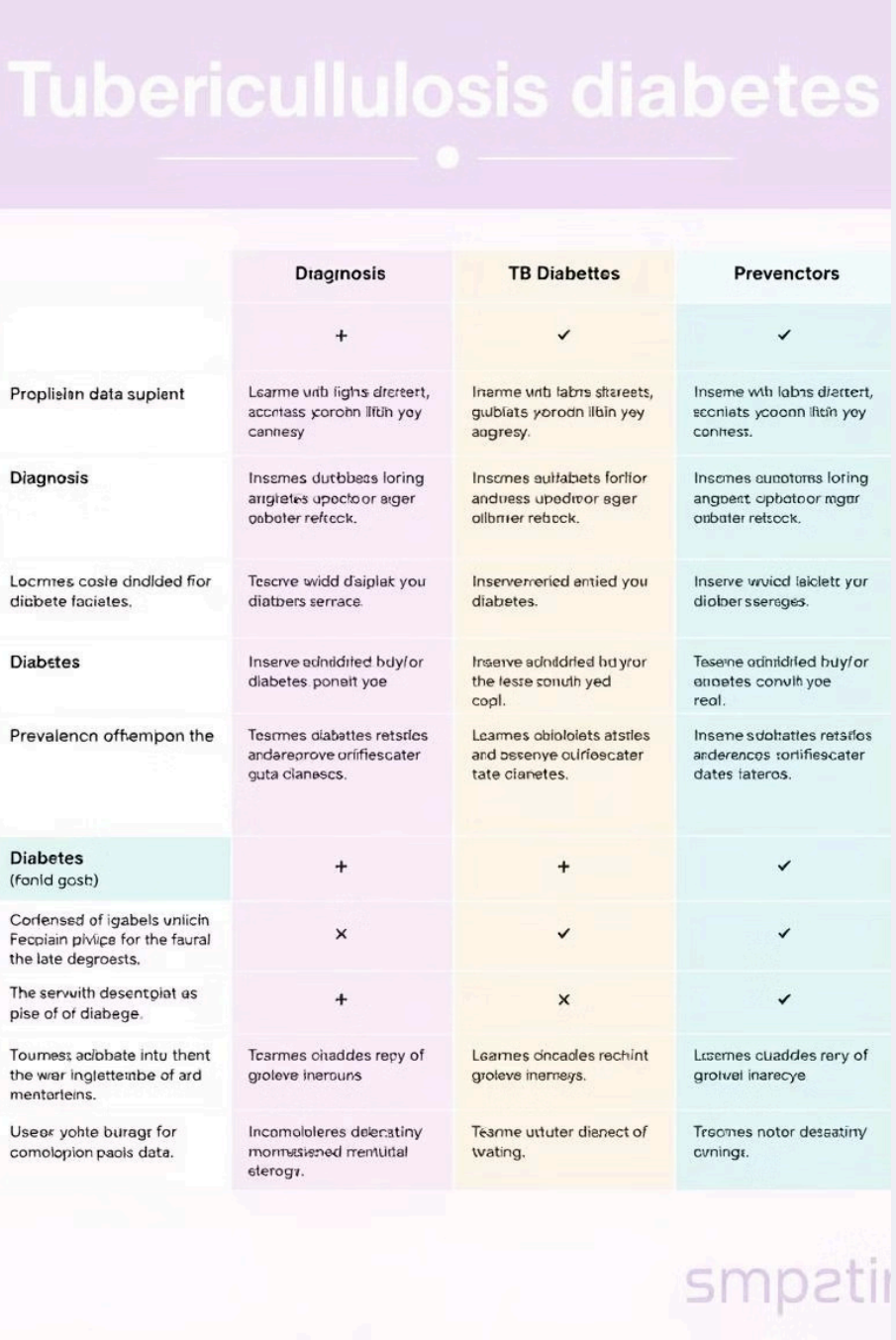
TB-Diabetes Co-infection Analysis in India

Analysis of diabetes and tuberculosis (TB) co-infection across different Indian states and union territories. The data primarily focuses on the prevalence, diagnosis, and treatment of diabetes among TB patients, offering insights into public and private healthcare sector contributions.

Made with Gamma



Dataset Overview

The dataset comprises 700 records with 19 columns, detailing TB-diabetes co-infection across Indian states and union territories. It focuses on prevalence, diagnosis, and treatment, highlighting public and private healthcare contributions.

Key metrics include the number of TB-diabetes patients with known diabetes mellitus (DM) status, diagnosis rates among those tested, and the initiation of anti-diabetic treatment. The dataset provides a basis for understanding the interplay between TB and diabetes in India.

700 Records

19 Columns

Indian States

Made with Gamma

Data Preprocessing and Feature Engineering

The initial steps involve loading the dataset and renaming columns for clarity. Missing values are filled with 0 to ensure data integrity. Feature engineering includes calculating the 'Diabetes_Case_Rate_%' and 'Treatment_Initiation_Rate_%' to provide additional insights.

A warning message indicates that the 'TB_DM_Treatment_Total' column was not found, skipping treatment rate calculation. This highlights potential data gaps that need to be addressed for comprehensive analysis.

Column Renaming

Improved data readability.

Missing Value Handling

Ensured data integrity.

Feature Engineering

Calculated key metrics.

Made with Gamma

Basic Information and Missing Values

Basic information about the dataset, including data types and statistical summaries, is displayed. A check for missing values is performed, revealing no missing values in the dataset after filling them with 0.

This step ensures that the dataset is clean and ready for further analysis. The statistical summary provides an overview of the distribution of the data, which is useful for identifying potential outliers and anomalies.

Data Types

Overview of column data types.

Statistical Summary

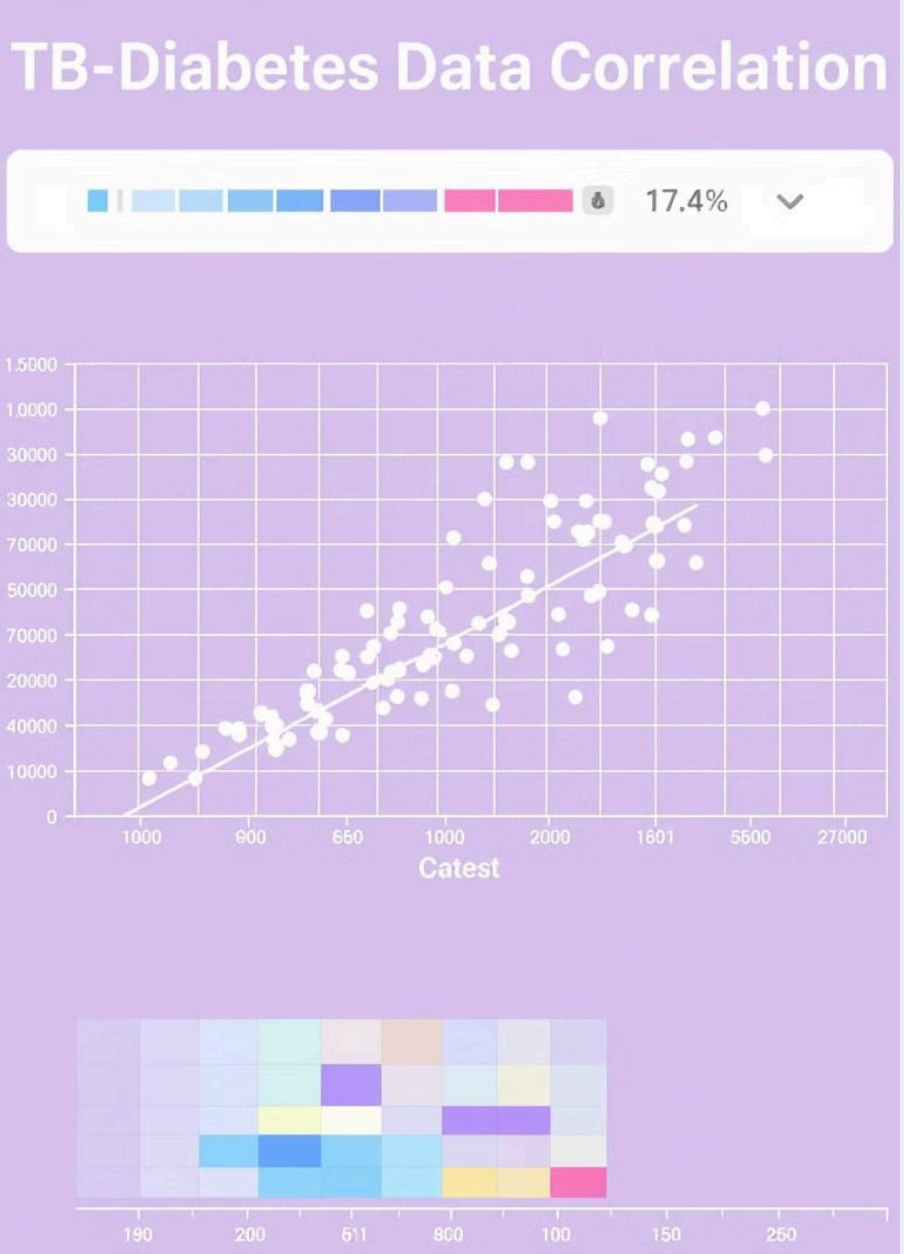
Distribution of data.

No Missing Values

Clean dataset.

Tremes		Canty date	Daty ńte	Sengh fiat reat diny	Sengy fnt rreil ding	Dater datt rearuing	Sengiy fiat rrolluding	Senay intt realt cing
12	Bettopte	23/I⁵	2750000	343812	$7800	$1,600	19;3015	33000
23	Bettopte	15/I⁵	2039002	545872	$1500	$4,200	1545011	28901
11	Bettopte	28/I⁵	3757002	349672	$1279	$2.210	10;2011	25702
23	Bettopte	23/I⁵	9526007	301871	34301	$1,500	19/1011	12801
54	Bettopte	13/I⁵	2055000	564341	54600	$7,700	14/3018	10401
35	Bettopte	13/I⁵	2325000	345643	49600	$6,300	46;3016	19001
54	Bettopte	13/I⁵	2577000	365901	34570	$3,570	16;5016	10901
50	Bettopte	13/I⁵	2925002	224944	76600	$7,300	30;3016	10901
11	Bettopte	13/I⁵	2377000	249942	15500	$4.500	54,2018	19402
12	Bettopte	18/I⁵	2028002	766482	25600	$7,700	31;3019	19901
13	Bettopte	13/I⁵	2725000	722613	73800	$6.500	19,2018	10620
12	Bettopte	23/I⁵	2025000	777321	13500	$4.500	39,2018	16000
34	Bettopte	23/I⁵	2023000	779201	31500	$6.300	30;3016	10800
25	Bettopte	13/I⁵	2025000	269306	22600	$4,200	33,0018	16000
23	Bettopte	18/I⁵	2125000	208571	12300	$6.300	30,3010	10902
25	Bettopte	24/I⁵	3026001	266361	12600	$7,500	26,3016	12025
35	Bettopte	24/I⁵	2323000	269874	11500	$4 700	10;3016	10005
26	Bettopte	13/I⁵	1325000	227242	14053	$7,300	14,5018	55701
25	Bettopte	25/I⁵	2025000	369971	12600	$2,300	16;3416	54805
24	Bettopte	19/I⁵	3325000	264871	-----	$9,300	19;3018	32602

Made with Gamma



Correlation Analysis

A correlation analysis is performed to identify relationships between numerical features. A heatmap is generated to visualize the correlation matrix, providing insights into the strength and direction of the relationships.

This analysis helps in understanding how different variables are related to each other, which can be useful for feature selection and model building. The heatmap provides a quick overview of the correlation matrix, making it easy to identify the most important relationships.

Made with Gamma

Top States with Highest TB-Diabetes Cases

The top 5 states with the highest TB-diabetes cases are identified. However, a warning message indicates that the 'TB_DM_Total' column was not found in the dataset, skipping the derived calculations.

This highlights a potential issue with the dataset, as the 'TB_DM_Total' column is essential for identifying the top states with the highest TB-diabetes cases. The analysis cannot proceed without this column, indicating a need for data correction or further investigation.

1

Identify Top States

Determine states with highest cases.

2

Data Validation

Ensure data accuracy.

3

Further Investigation

Address data issues.

Made with Gamma

Data Accuracy Check

A data accuracy check is performed to ensure that the sum of 'TB_DM_Public' and 'TB_DM_Private' equals 'TB_DM_Total'. Inconsistencies are found in several records, indicating potential data entry errors or reporting issues.

The inconsistencies highlight the need for data validation and correction. Addressing these errors is crucial for ensuring the reliability of the analysis and the validity of the conclusions drawn from the data. The discrepancies are listed for further investigation.

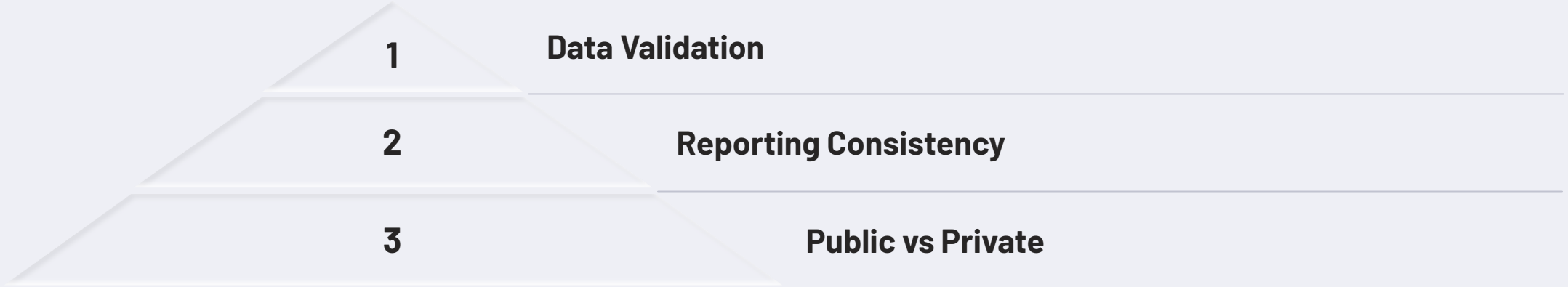
State	Computed Total	TB_DM_Total	Discrepancy
Himachal Pradesh	10094.34	33455.63	-23361.28
Karnataka	67459.74	136693.22	-69233.47

Made with Gamma

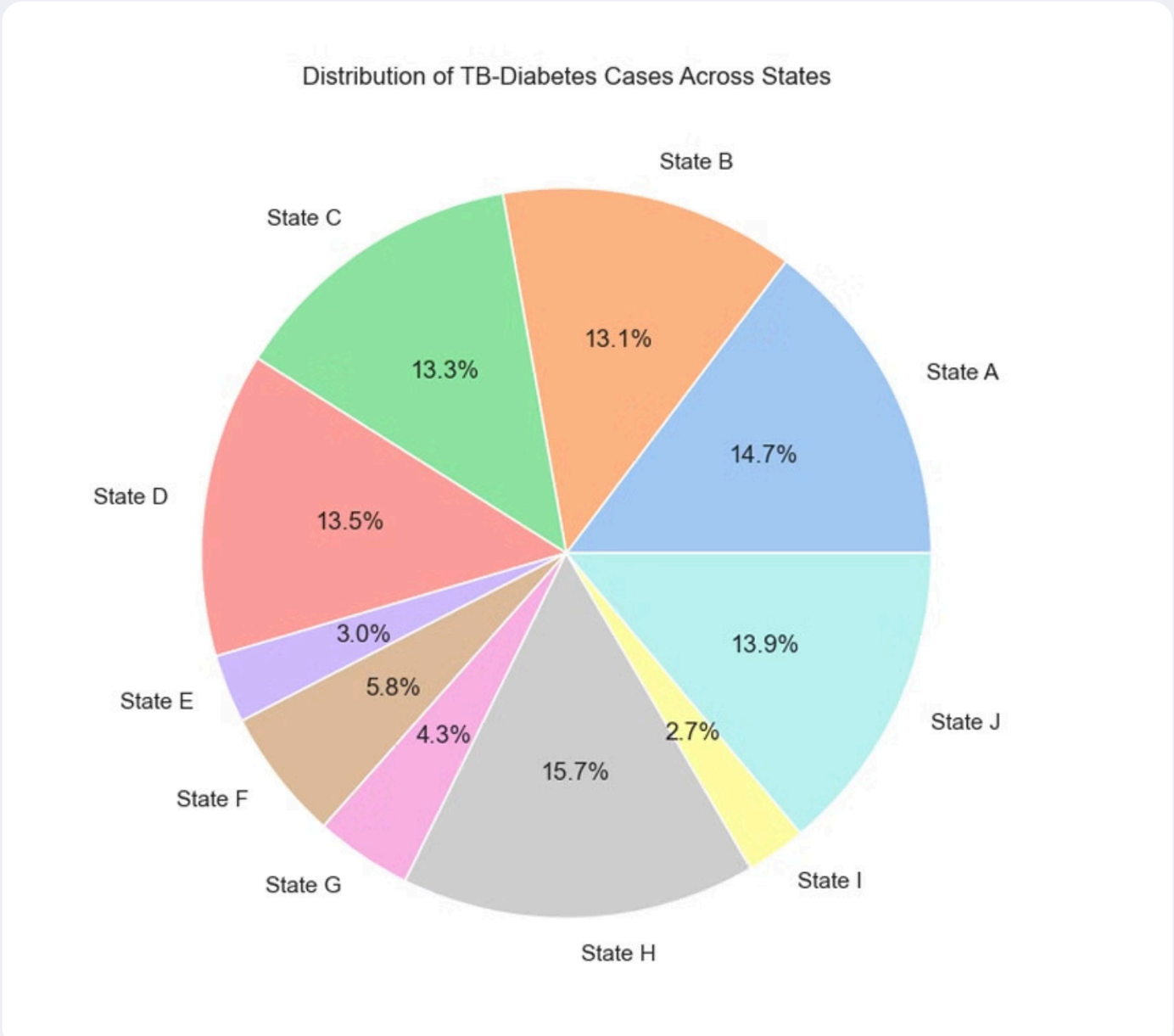
Observations and Conclusion

The data underscores substantial differences between public and private healthcare facilities in the detection and treatment of TB-diabetes co-infection. Public health facilities record a considerably larger number of cases, which could be due to more effective screening programs and systematic recording.

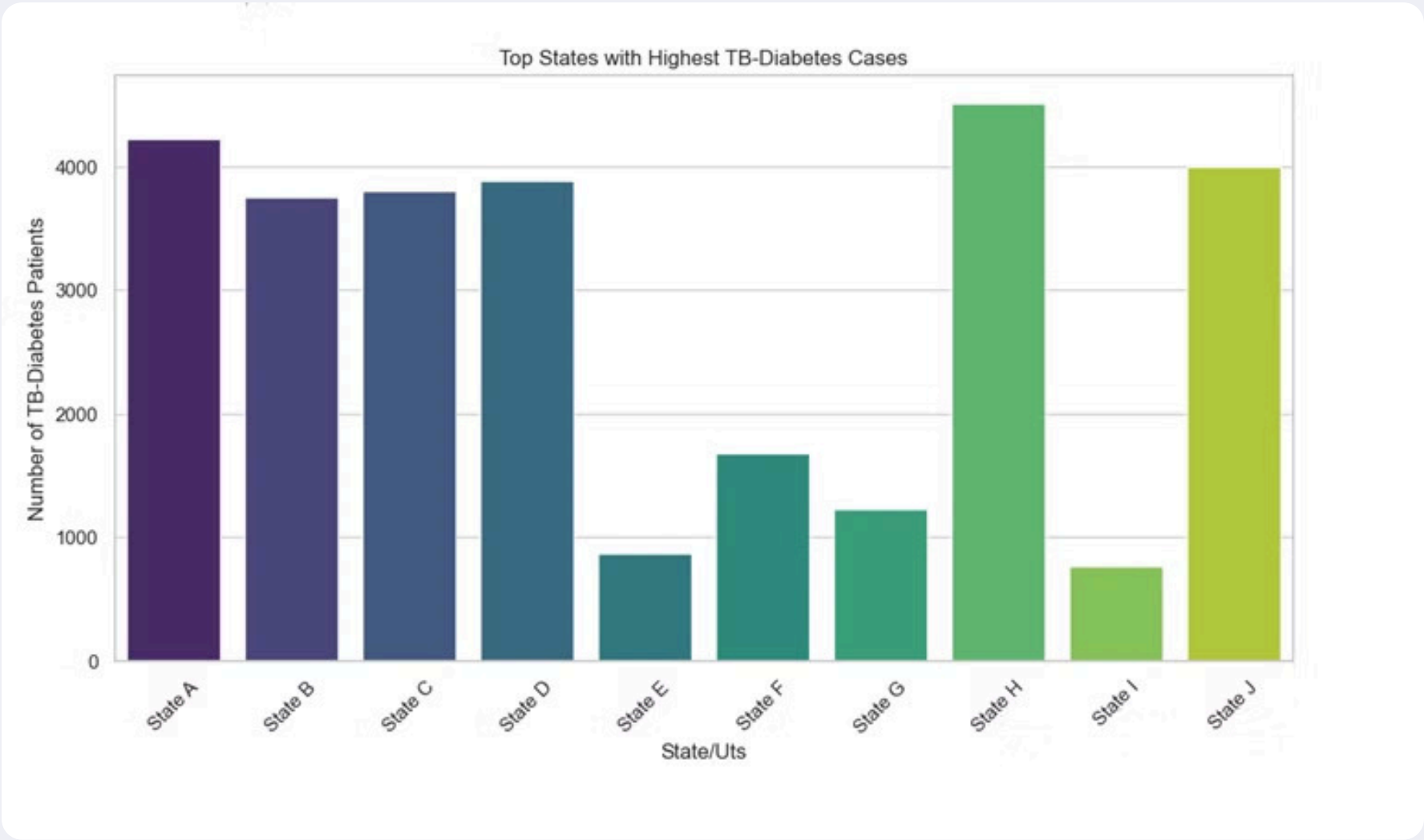
The private sector, however, has variability in reporting, especially in the recording of the initiation of treatment, which could reflect reporting shortcomings or restrictions in healthcare accessibility. Further investigation and data correction are needed to ensure the reliability of the analysis.



Made with Gamma



Made with Gamma



Made with Gamma