# INSURANCE DATA ANALYSIS

➢ Here are some points to be noted before moving forward.

- All the graphs are generated based on CSV file where data is randomly generated. Data is fully synthetic.

- You may get to see there is almost equal sharing of data in several Sales channel, Policy type, Agent types, Financial years etc.

- Data has been shared you may make changes according to you and plot graphs.

- All the codes are also shared which will generate the graphs and you may also alter the code once you download the code and CSV file.

- This PPT is just to represent the outcomes of code and do not contain extensive explanation. Only mentioned some points to give explanation superficially.
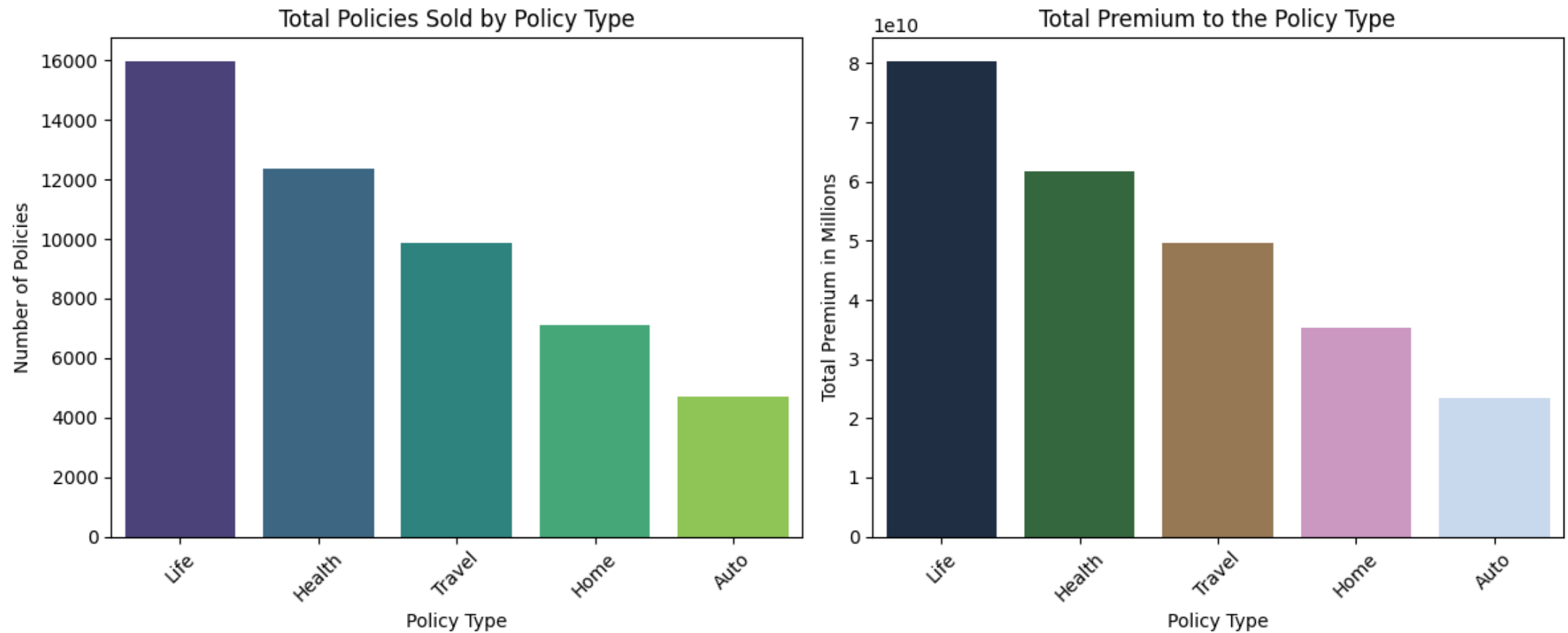
❖ This graph shows the summary of premium collected and number of policies sold out in all the financial years.

  FY 2024-25 is most beneficial year. Also it shows the consistent growth of Company.
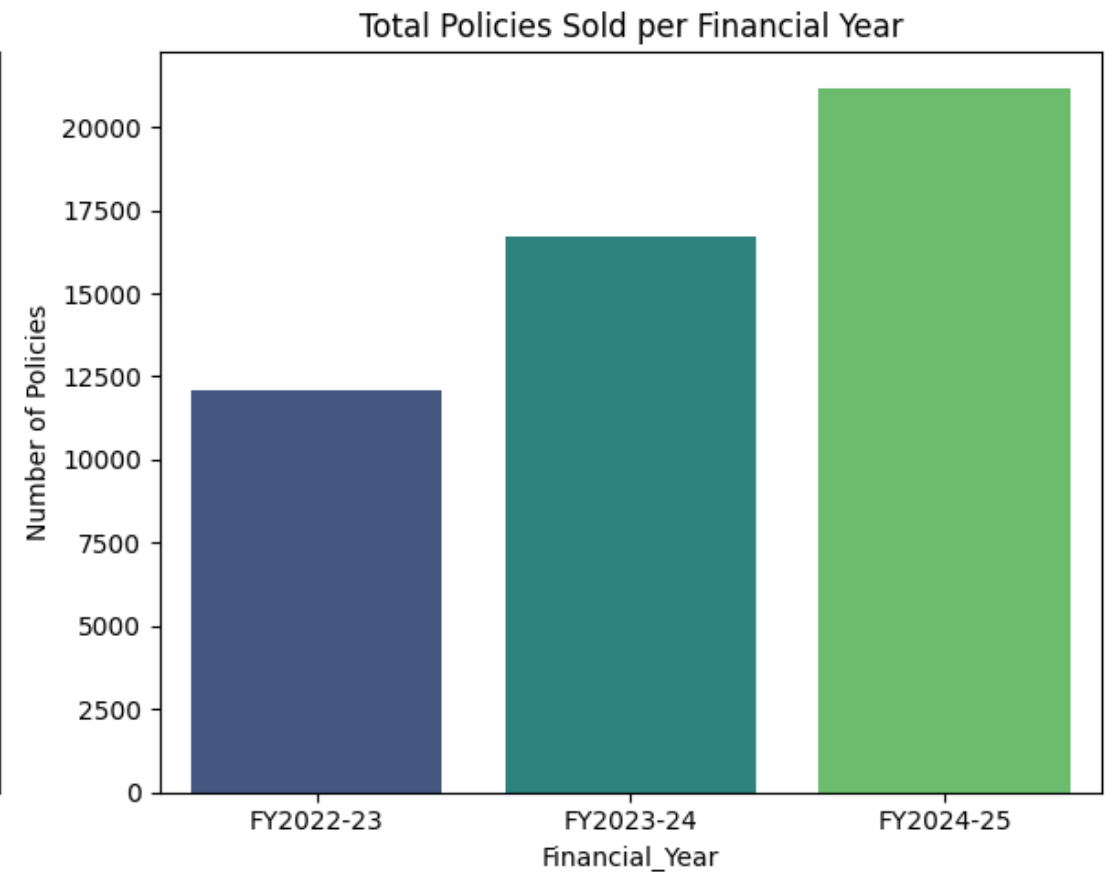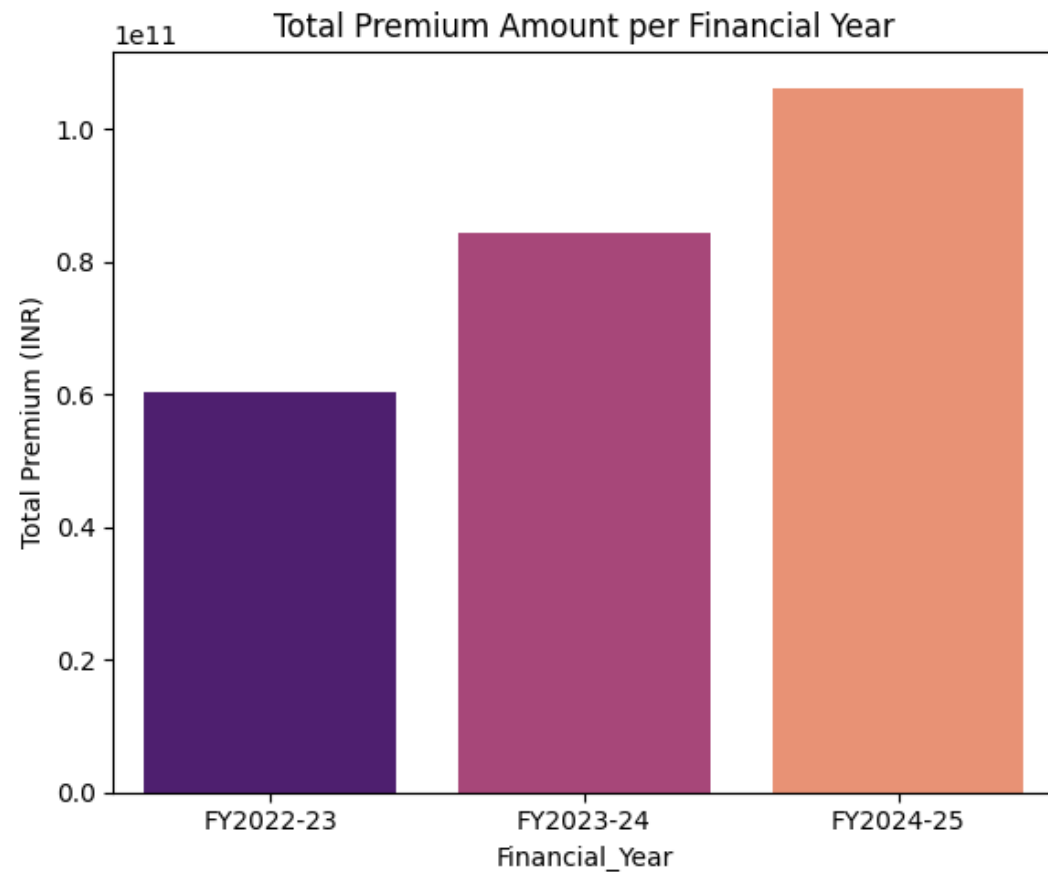
❖ This graph shows the summary of premium collected and number of policies sold out in various policy types.

Life is mostly sold followed by health. Auto is least preferred.

❖ This pie chart shows the percentage distribution of all the policy type sold in all the financial years.

Life contributes most 32% followed by Health whereas Auto has least shares i.e., 9.4% among all policy types.


Percentage of Total Premium by Policy Type
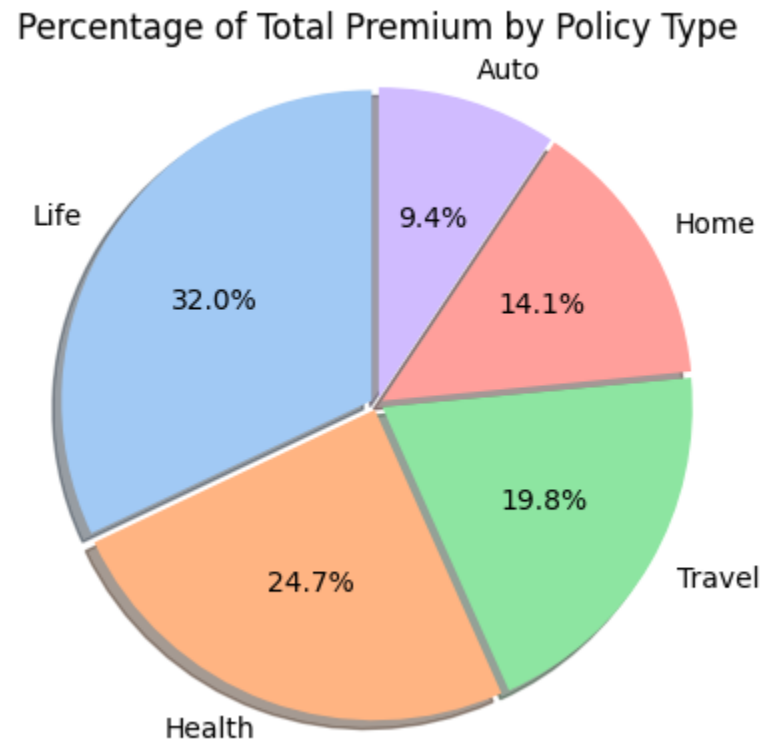
❖ This graph shows that the company has a large and fairly evenly distributed customer base across a wide age range (approximately 20 to 85 years), with the highest concentrations appearing in specific, but not dramatically distinct, age groups within that range.

Distribution of Customer Age
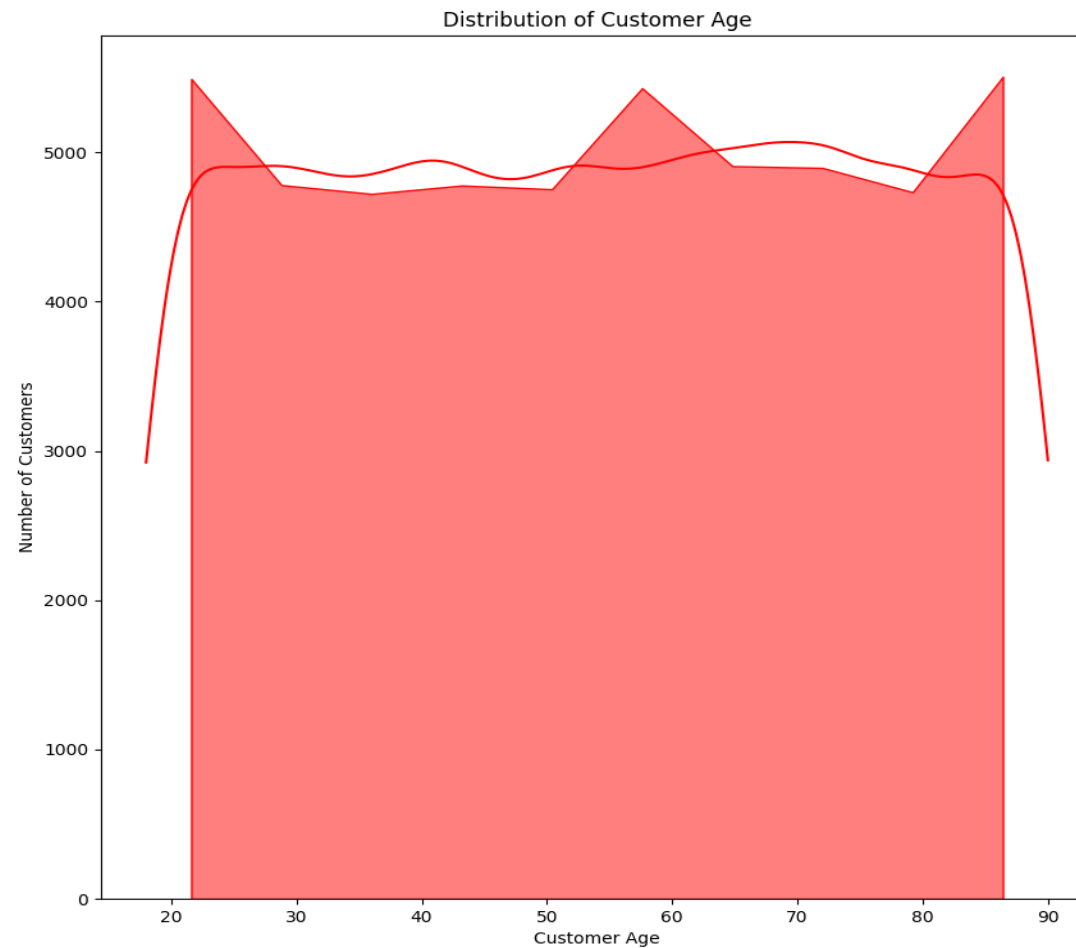
❏ This graph shows 2 details

1- Total premium paid by several age groups. <mark>65-74 and 75+ contributes most.</mark>

2- Total premium paid be all the genders. <mark>Male and females are almost same.</mark>

Premium paid by Age Group


Premium paid by diffrent Genders
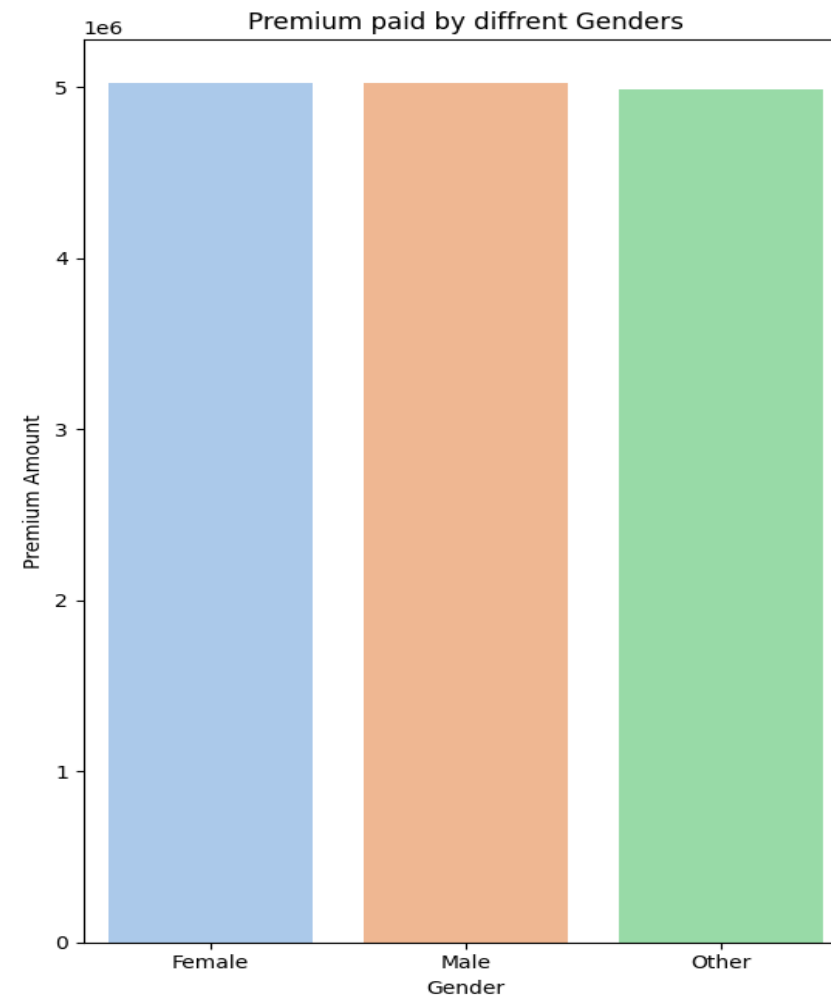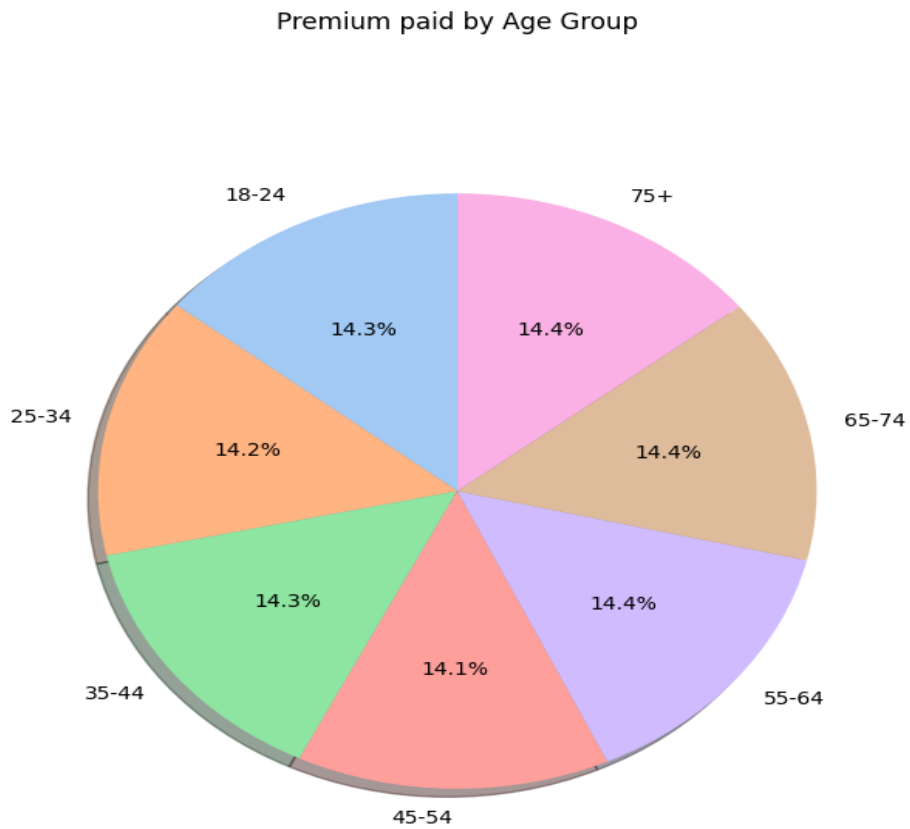
❖ Rajasthan is the 1st in top 10 countries in selling the policies and Odisha is 10th .

(You can change the colour of the graph. Palette's set is given in Main.py file you can use it, I just randomly give red colour 😁 .)

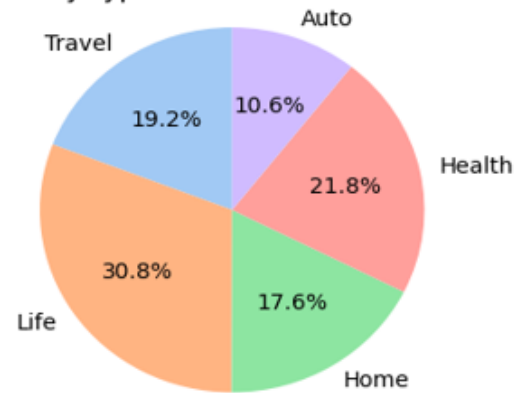Top 10 States by Number of Policies Sold.
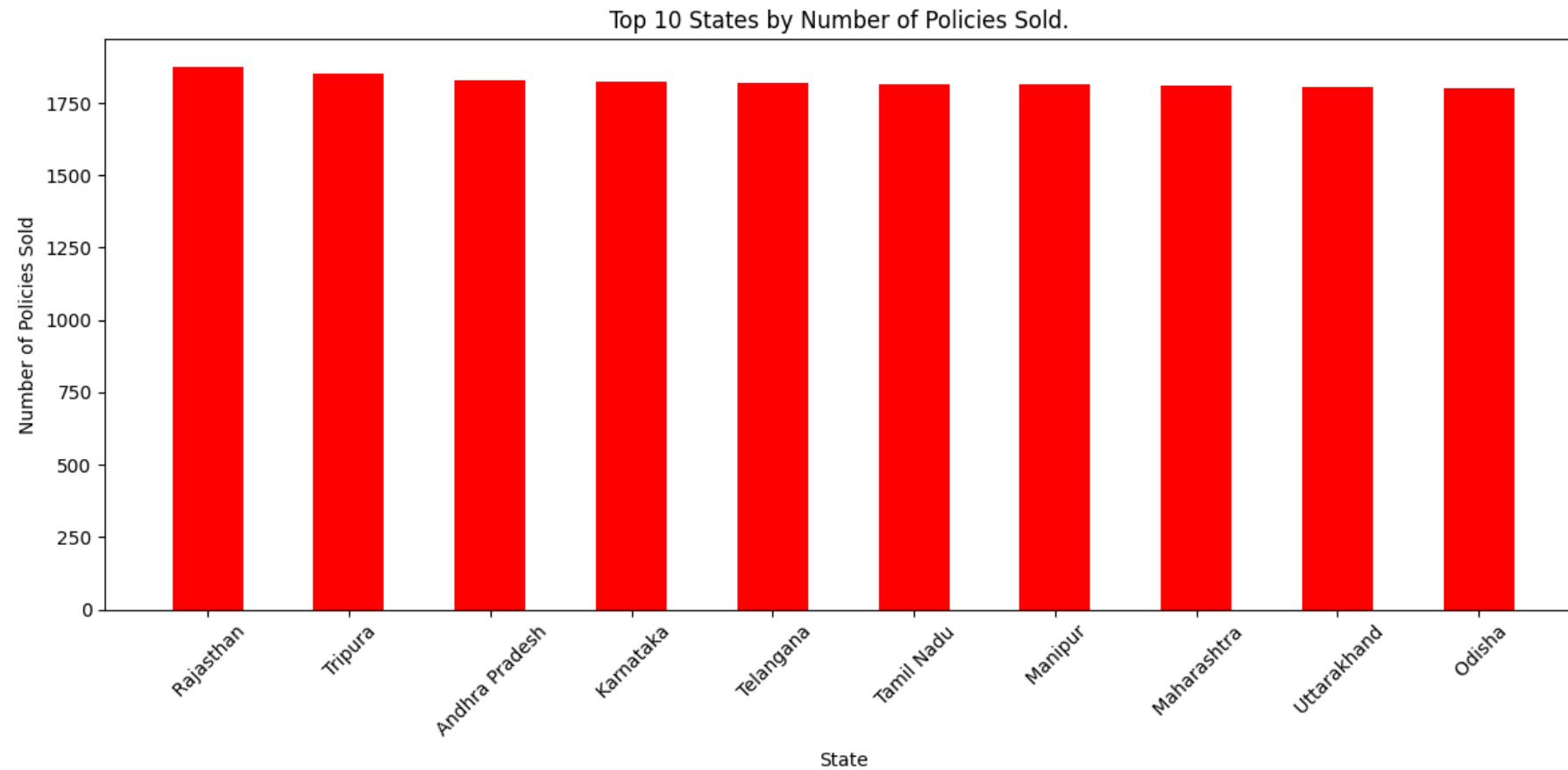
❖ Agents sold most of the policies and also generated the largest business followed by Brokers.
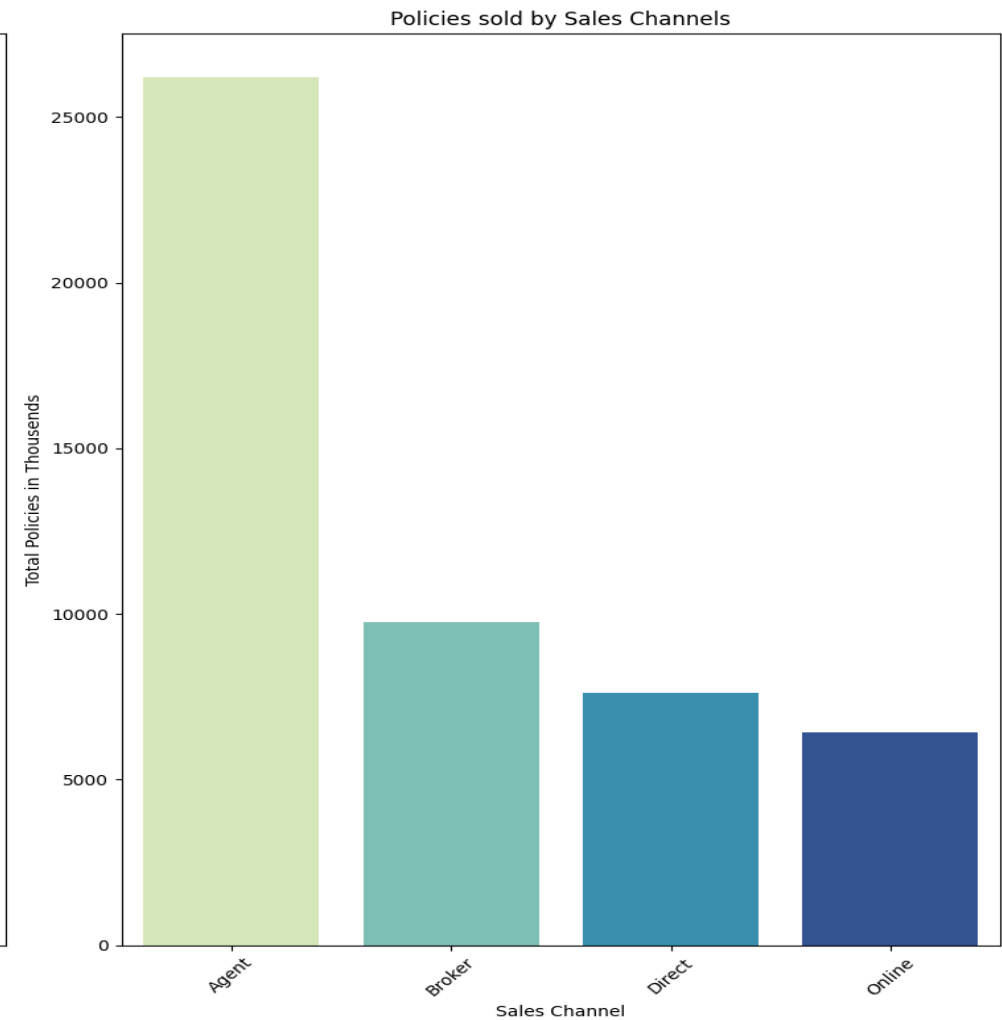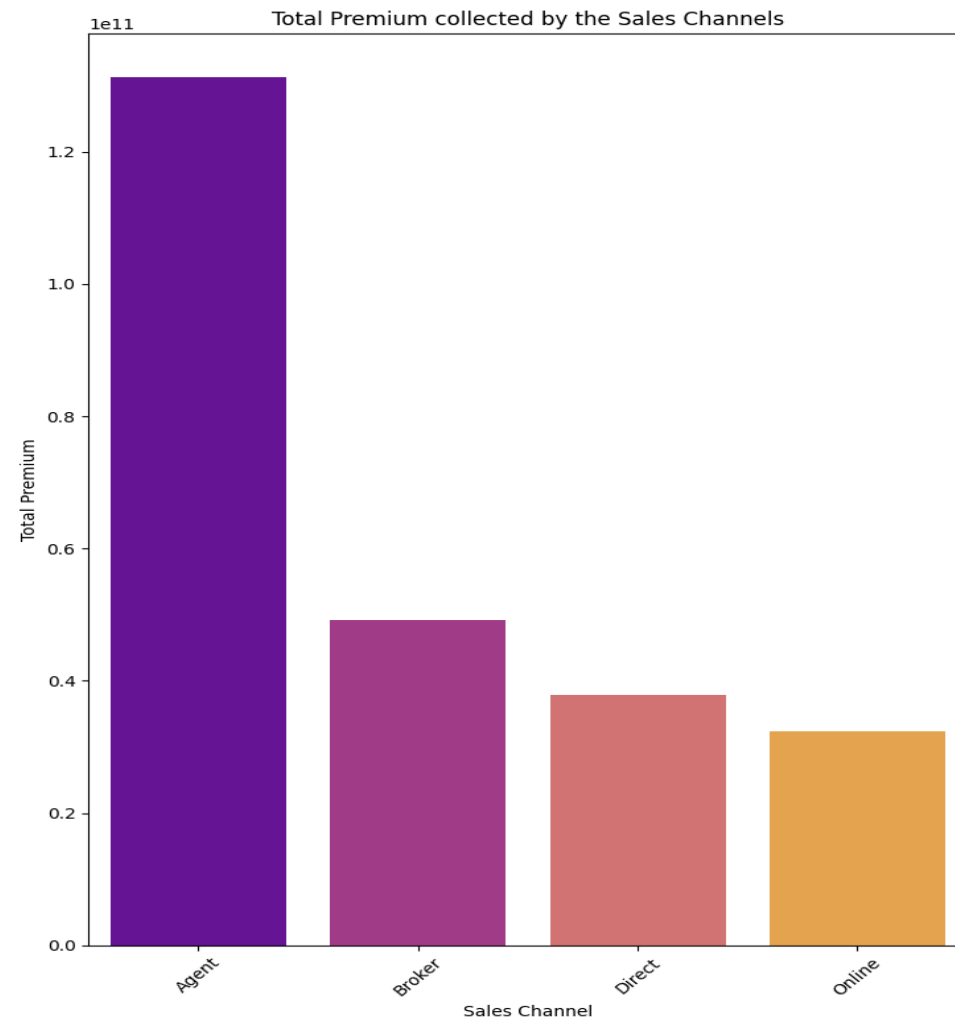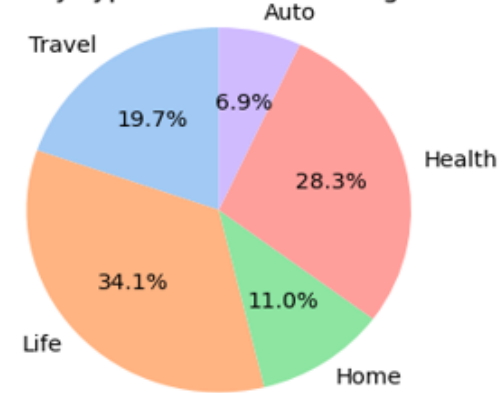
❖ Types of max policies sold by various sales channels

All of the sales channels sold Life insurance policies followed by health. Auto insurance is least shared among all others.
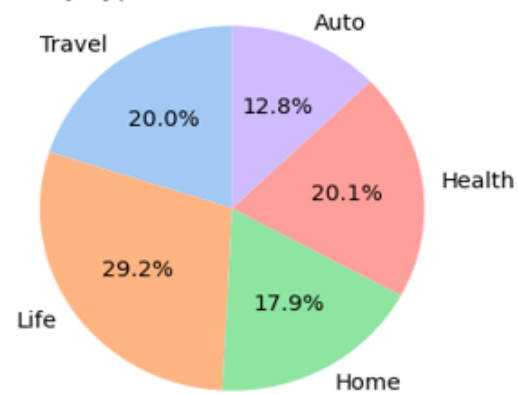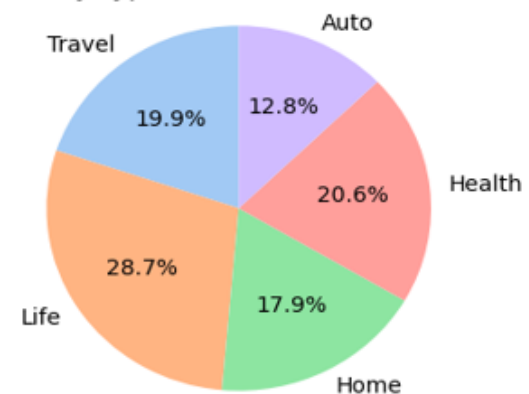

Policy Type Distribution for Online

Auto 10.6%
Health 21.8%
Home 17.6%
Life 30.8%
Travel 19.2%

Policy Type Distribution for Agent

Auto 6.9%
Health 28.3%
Home 11.0%
Life 34.1%
Travel 19.7%

Policy Type Distribution for Broker

Auto 12.8%
Health 20.1%
Home 17.9%
Life 29.2%
Travel 20.0%

Policy Type Distribution for Direct

Auto 12.8%
Health 20.6%
Home 17.9%
Life 28.7%
Travel 19.9%
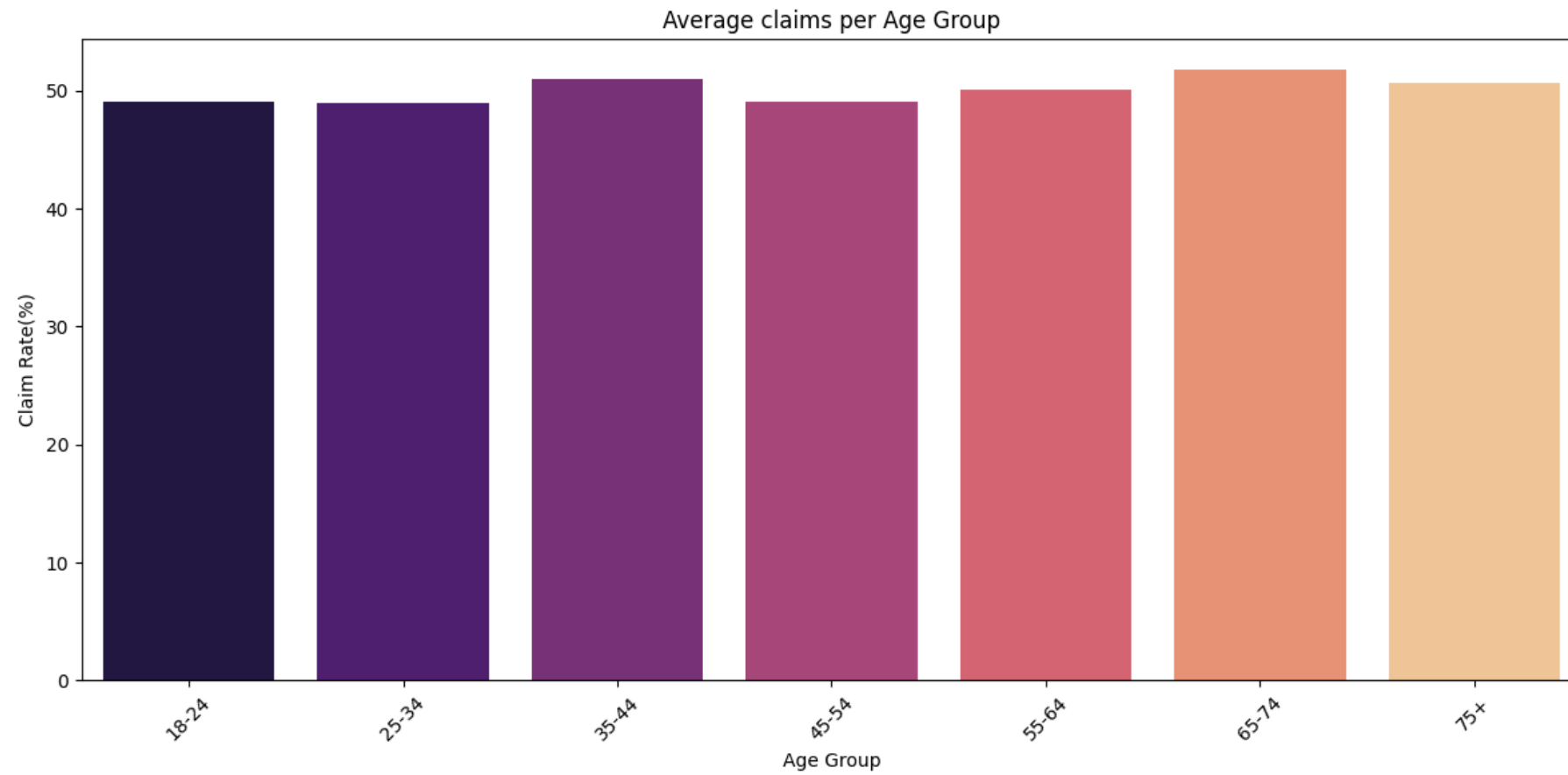
❖ Average claim from all the policy type is almost same , although Health Policies have higher claims as compared to others.

Average clams per Policy Type
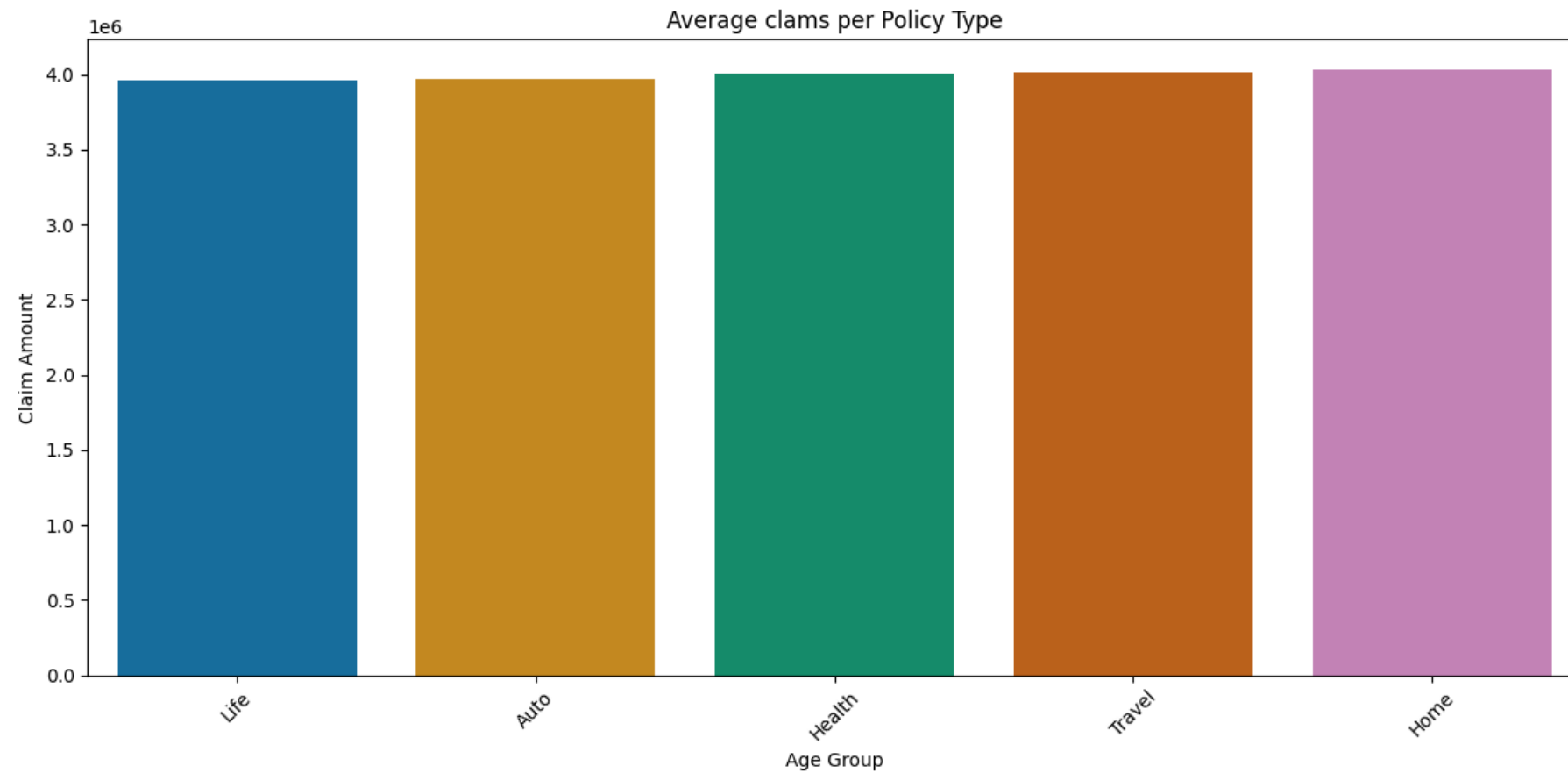
❖ Age group 65-74 has claimed most followed by 35-44.
   Least claims are from 25-34 age group.

Average claims per Age Group

❖ Assam has lowest number of claims whereas Gujarat is more prone to claims and has the highest number of claims.

Claim Amount of all the States
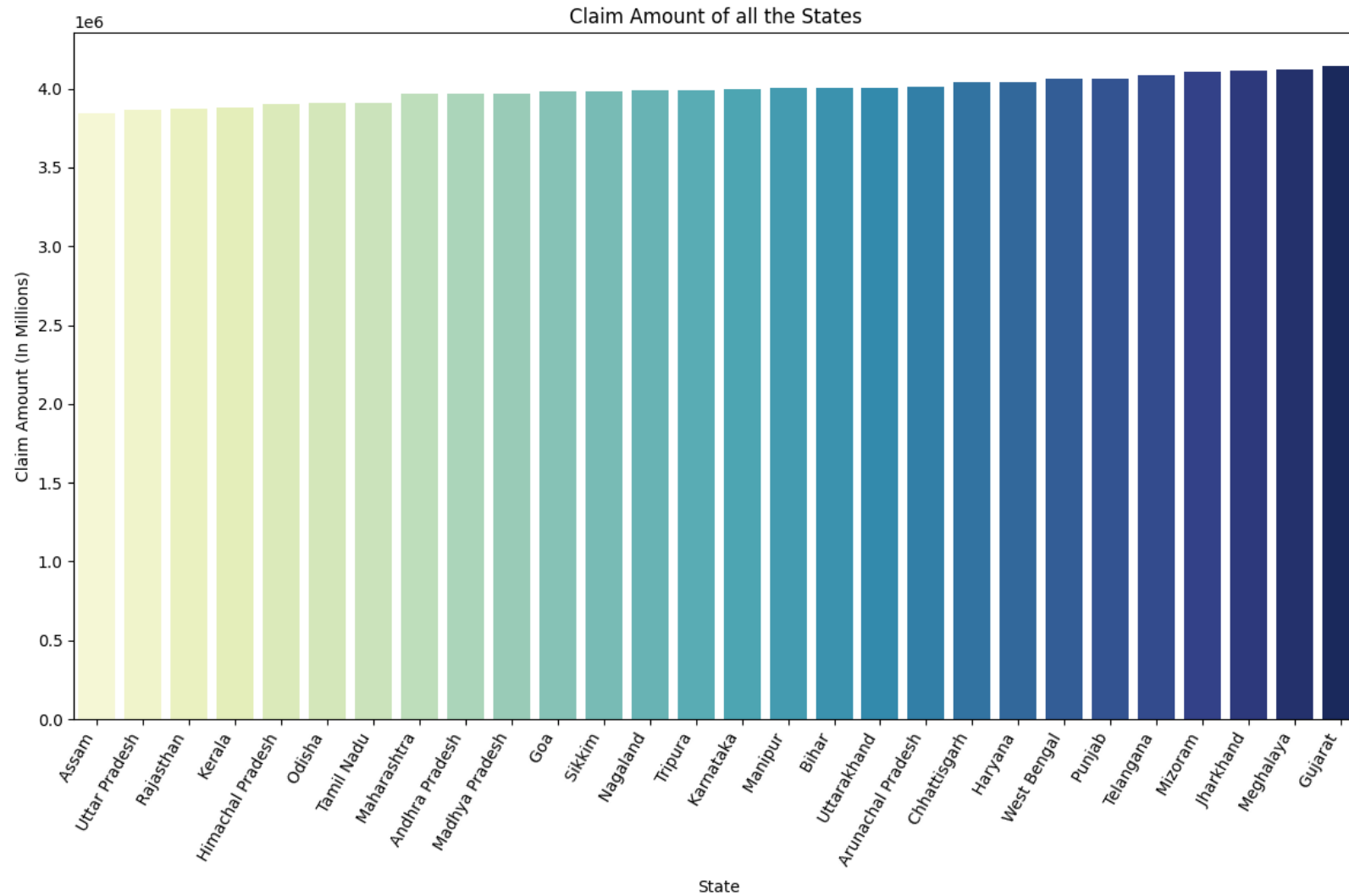
❖ Sikkim is lowest state in generating the Premium but Himachal Pradesh is highest in generating revenue


Premium Amount of all the States
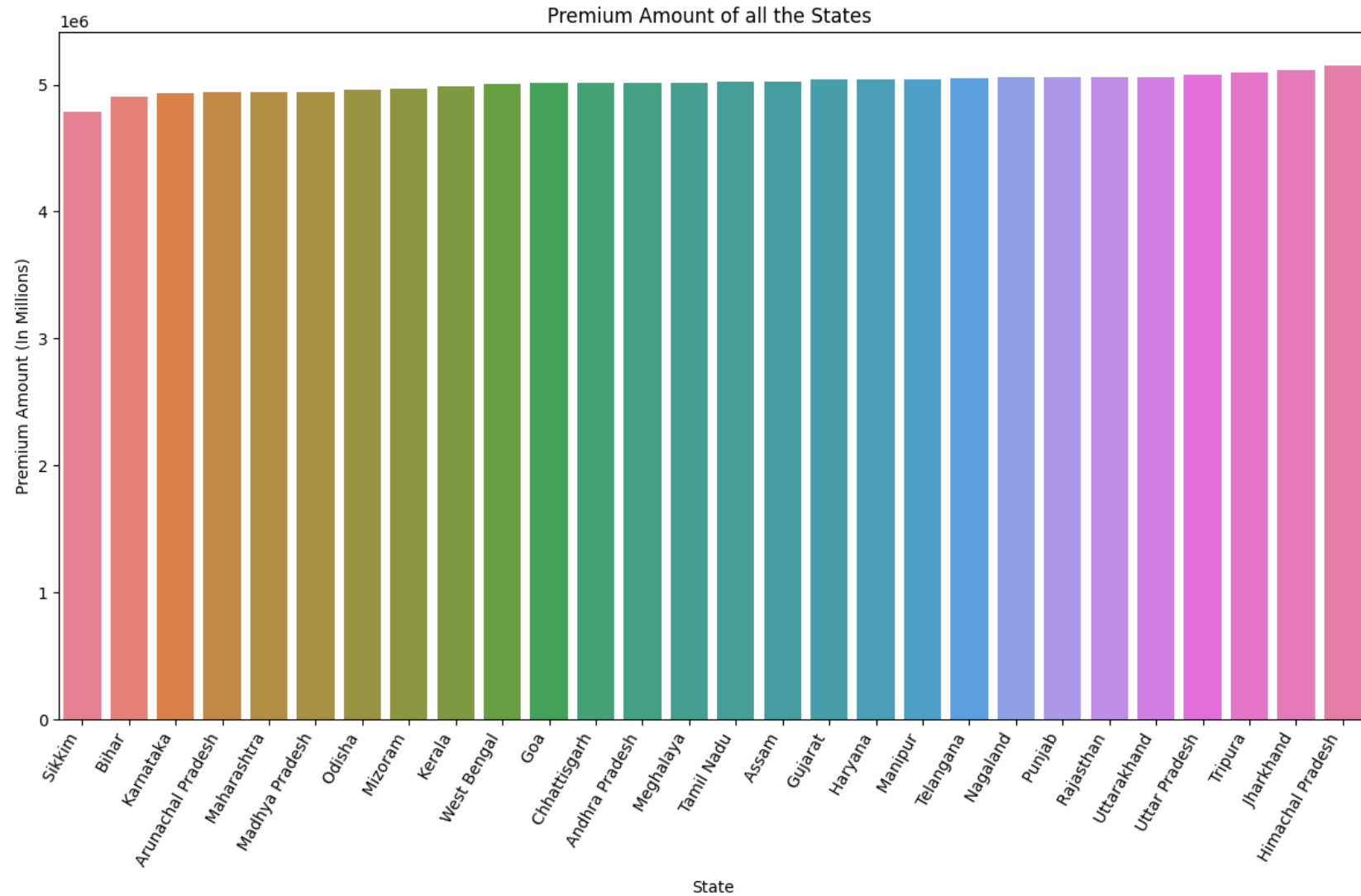
❖ Premium to claim ration of all the state where Himachal Pradesh has highest and Maharashtra has lowest which indicates HP has high Premium compared to Claim

Premium to claim ratio vs. State
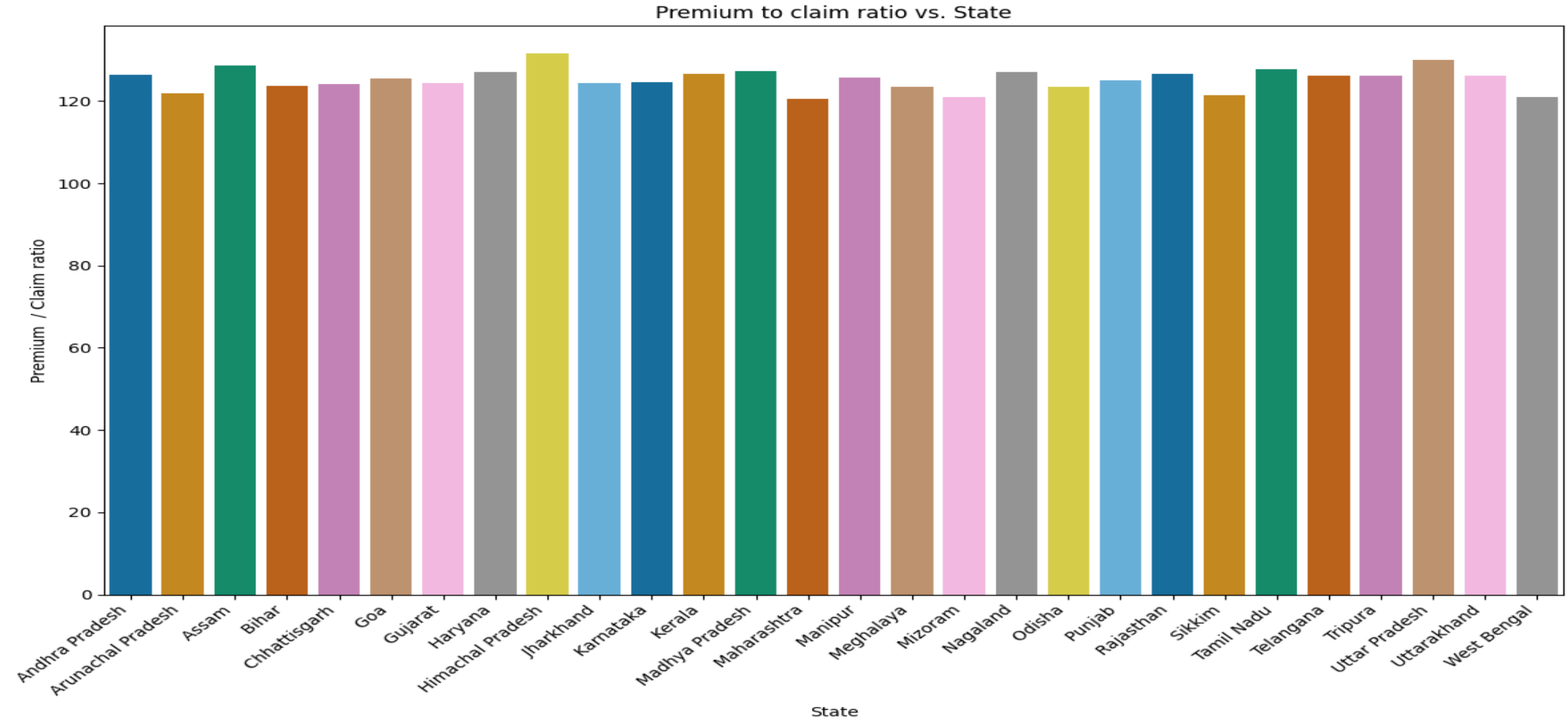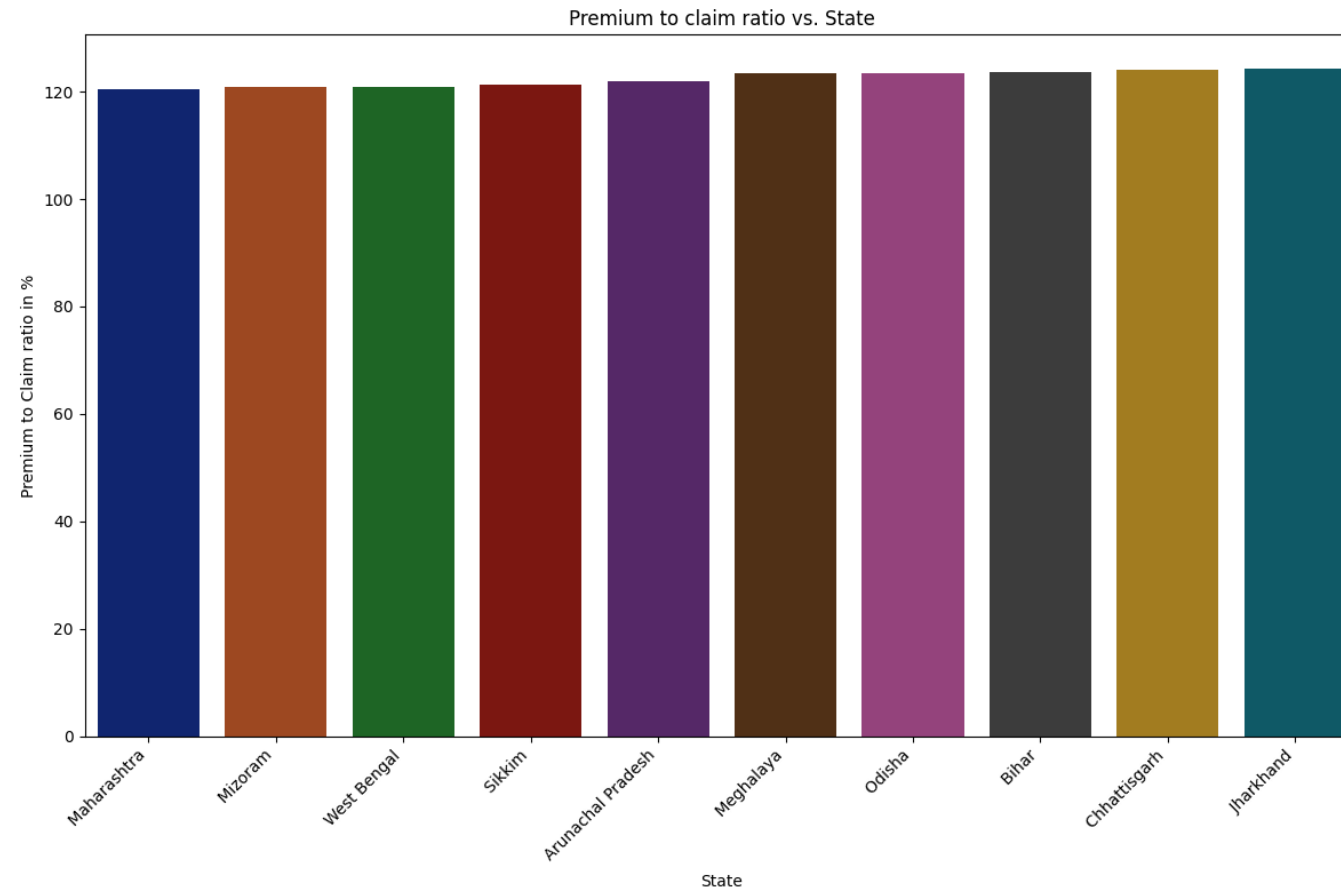
❖ Top 10 states which have high claim ratio.

Himachal Pradesh has highest claim ratio and Andhra Pradesh has lowest among top 10.
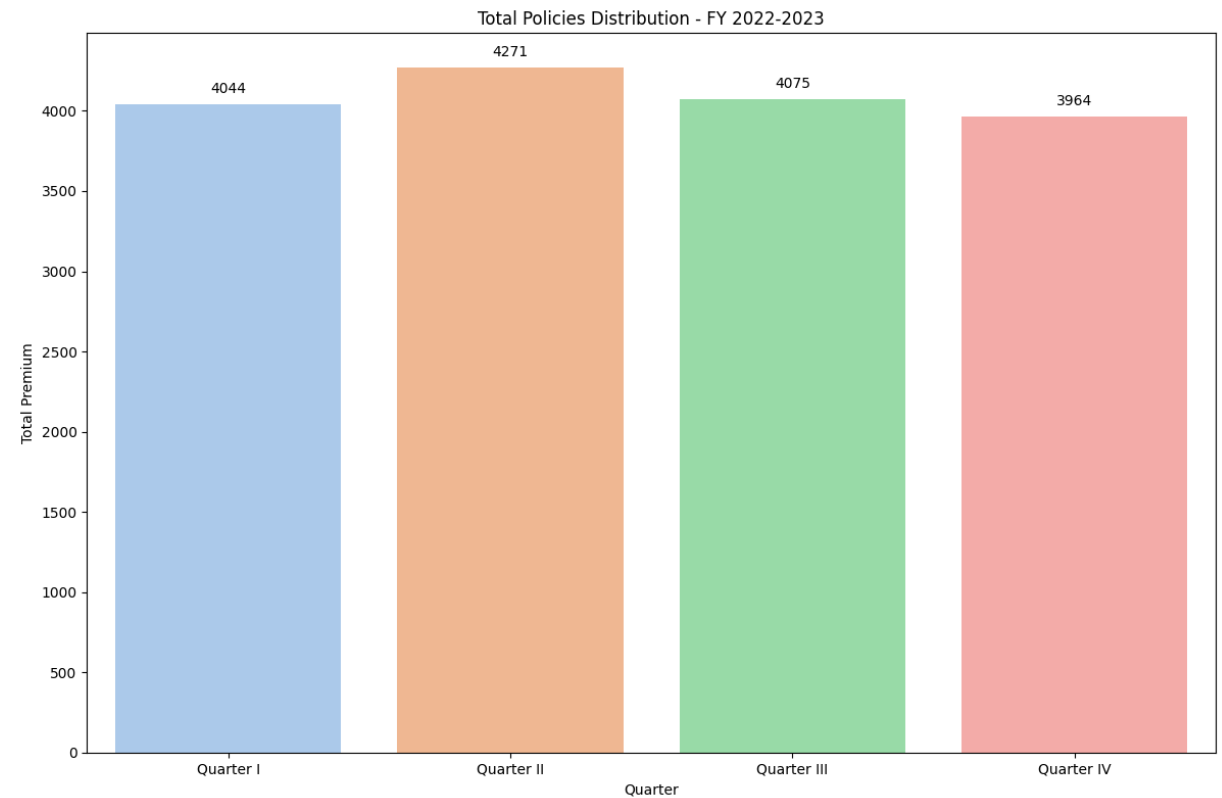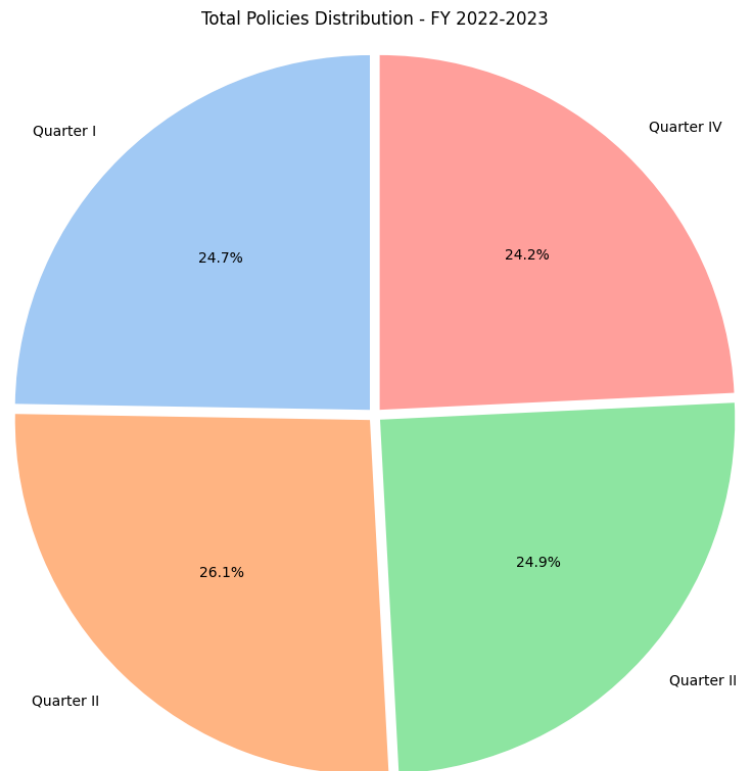

Premium to claim ratio vs. State
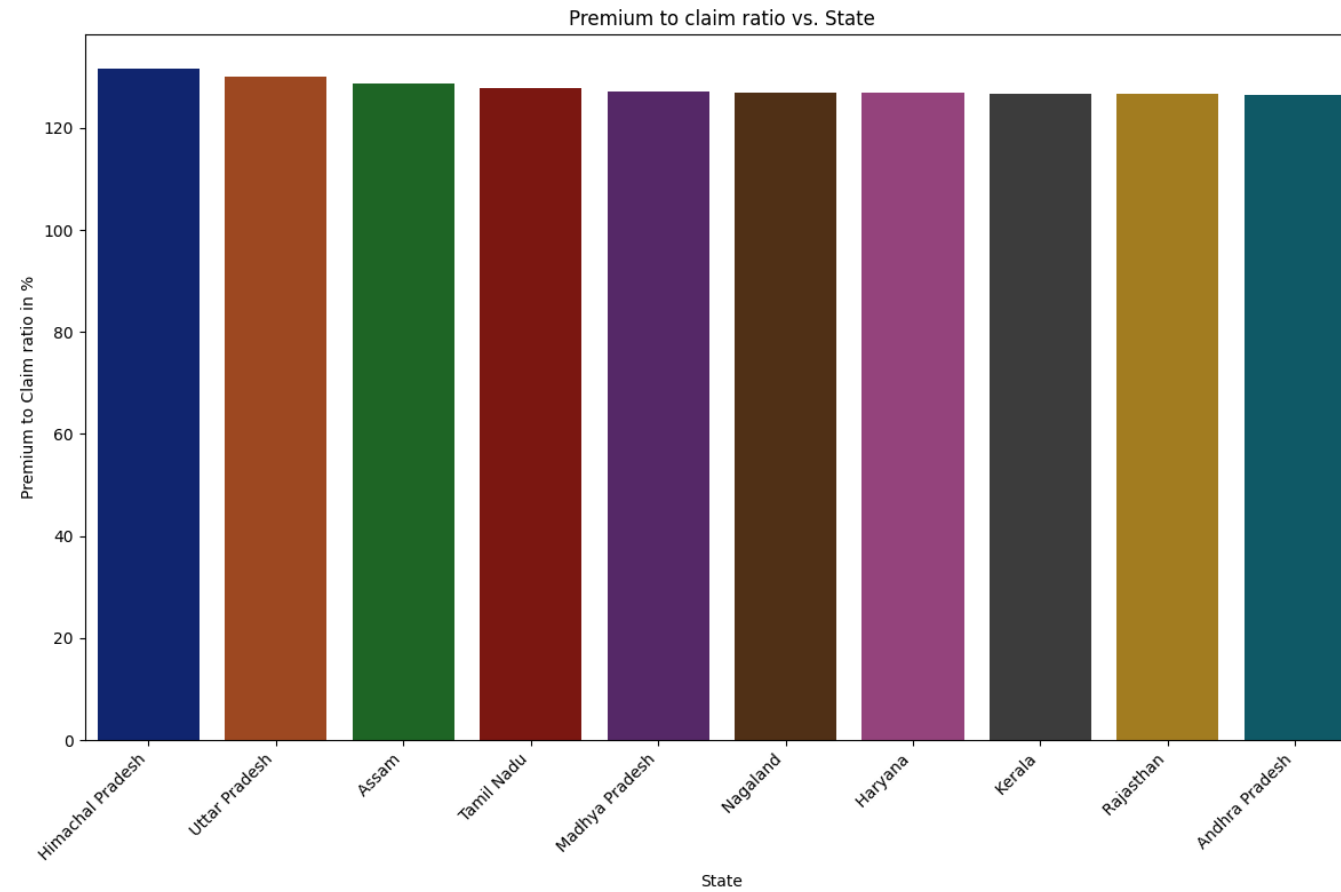
❖ Top 10 states with low claim ratio.

Maharashtra has lowest claim ratio whereas Jharkhand has highest in top 10.

❖ Quarter 2 of FY 2.22-23 has highest %age of policy sales having count 4271 followed by Quarter 3 (4075)

Total Policies Distribution - FY 2022-2023

Total Policies Distribution - FY 2022-2023

❖ Quarter 2 of FY 2023-24 has highest %age of policy sales having count 4295 followed by Quarter 1 (4188)

Total Policies Distribution - FY 2023-2024

Total Policies Distribution - FY 2023-2024
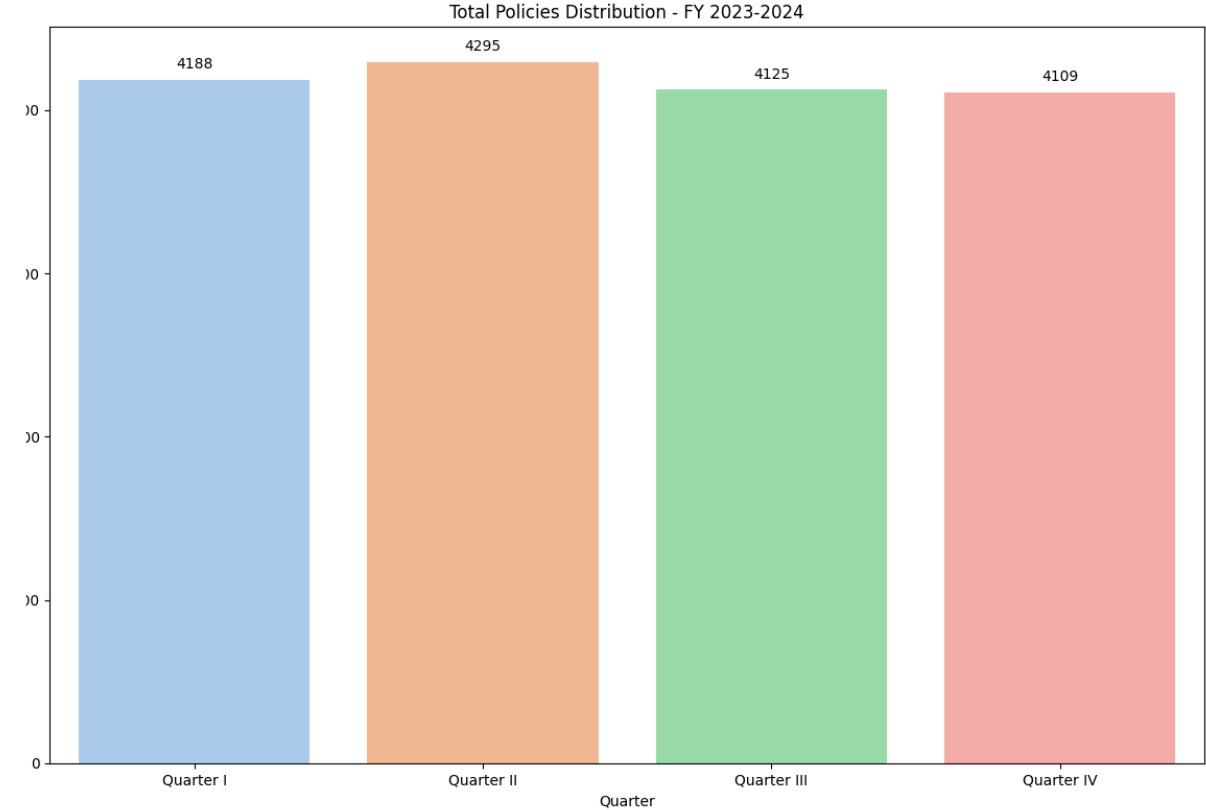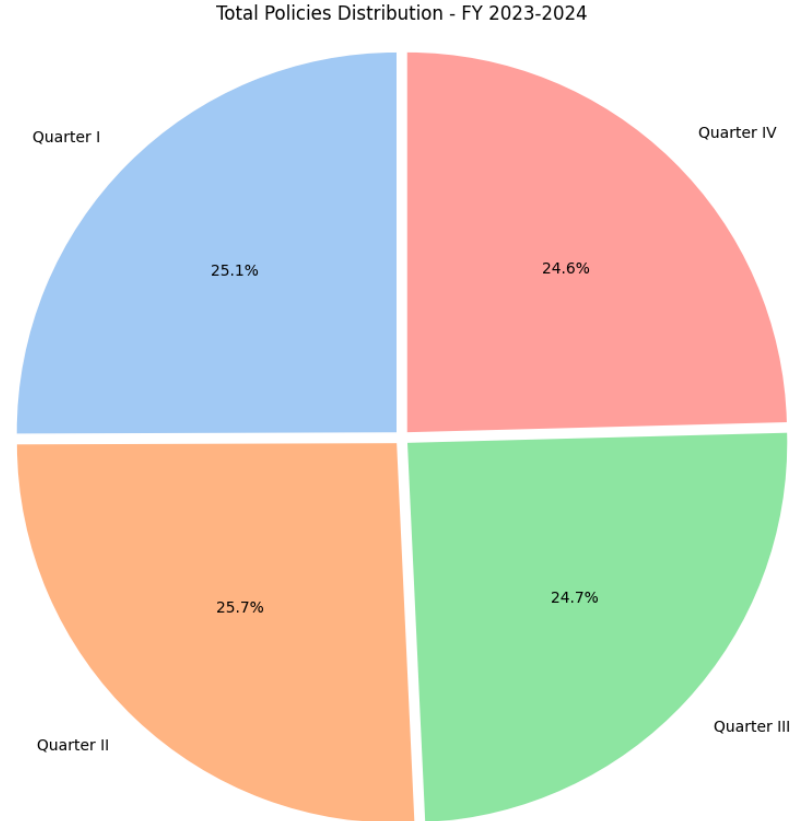
❖ Quarter 2 has highest %age of policy sales having count 4372 followed by Quarter 3 (4310)

Total Policies Distribution - FY 2024-2025

Total Policies Distribution - FY 2024-2025
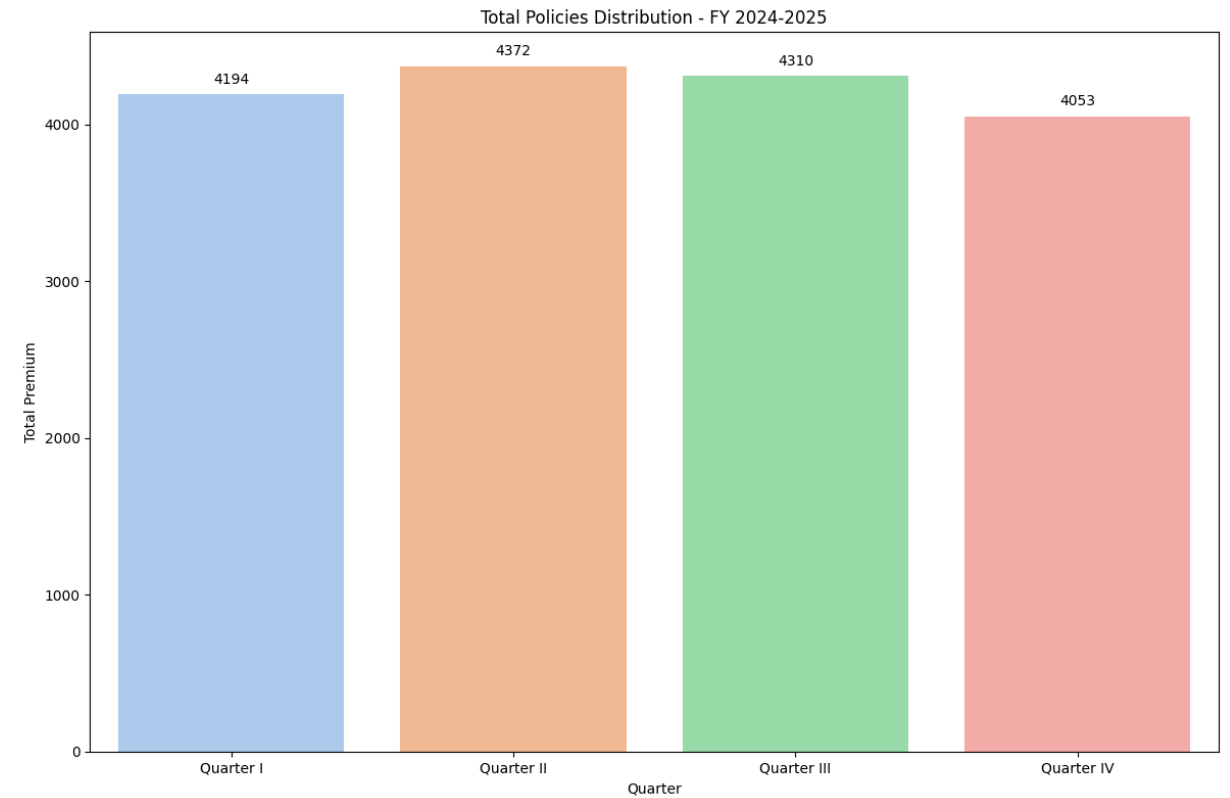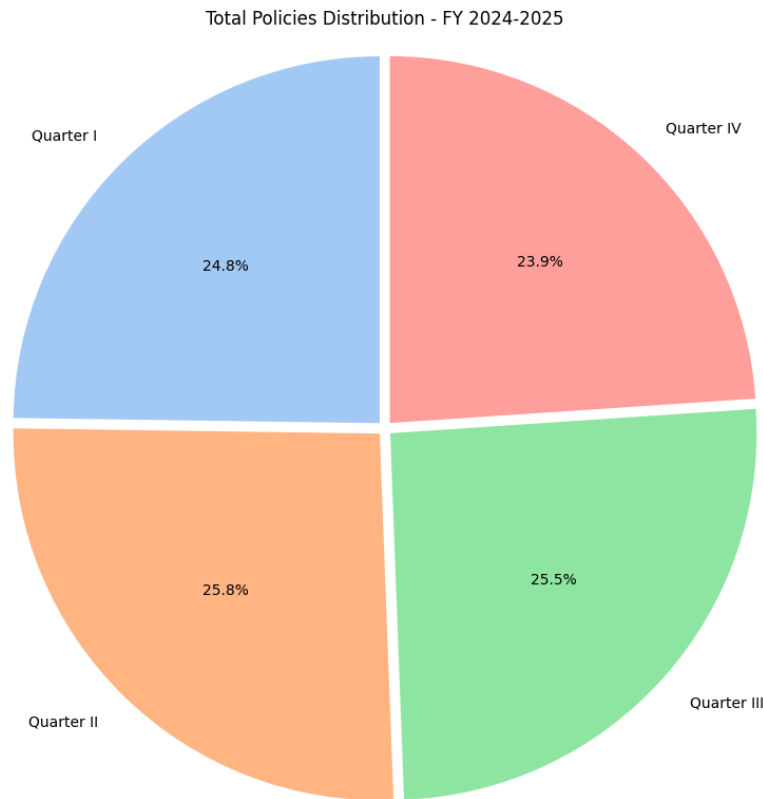
❖ FY 2024-25 was more profitable but sales dropped in Q4 indicating business loss.

Quarter-wise Policy Sales Trend showing Premium collection
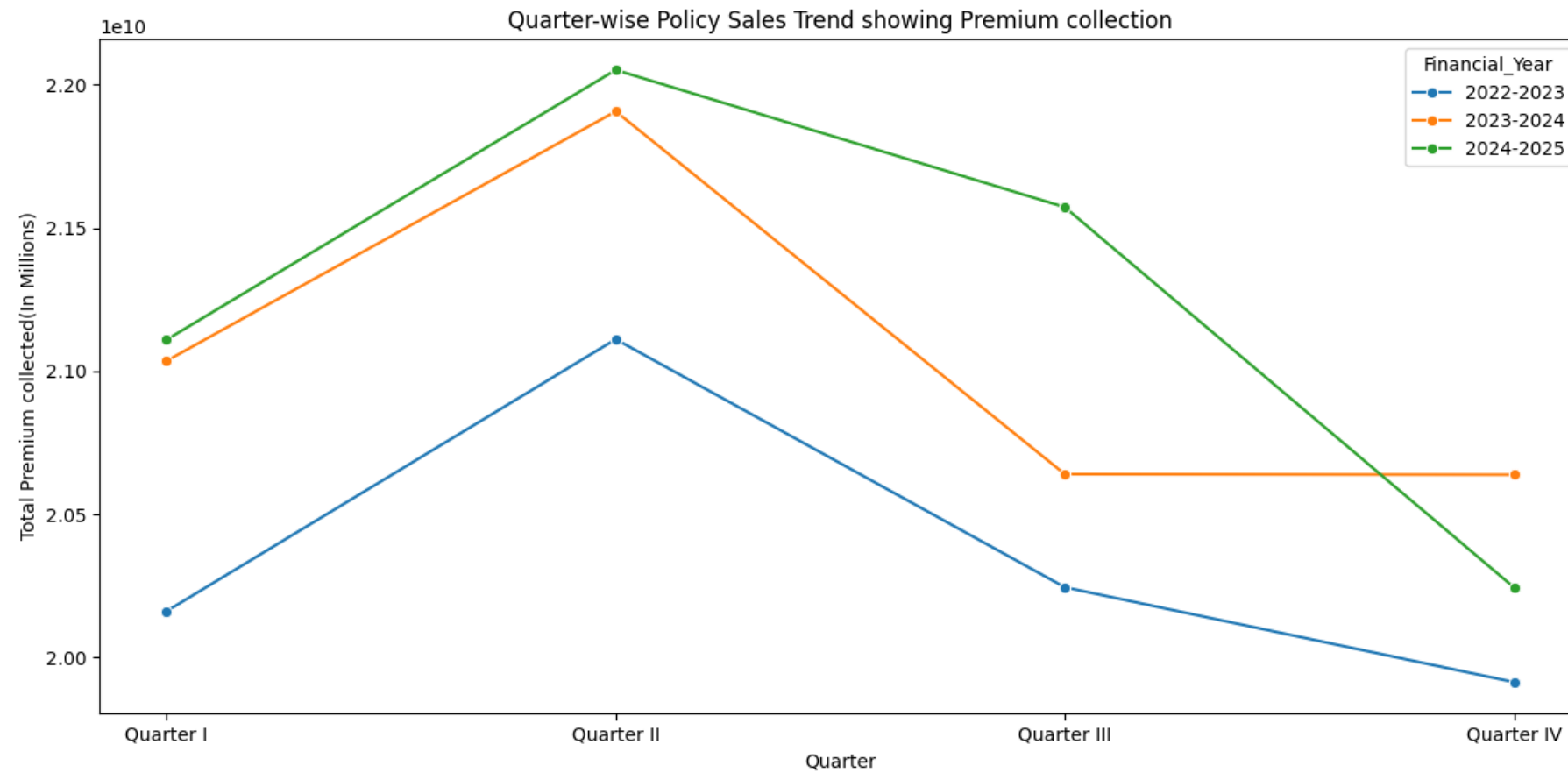
❖ Q3 is most profitable among all the financial years. In FY 2024-25 Q4 show less sales then Q4 of FY 2023-24.

Quarter-wise Policy Sales Trend
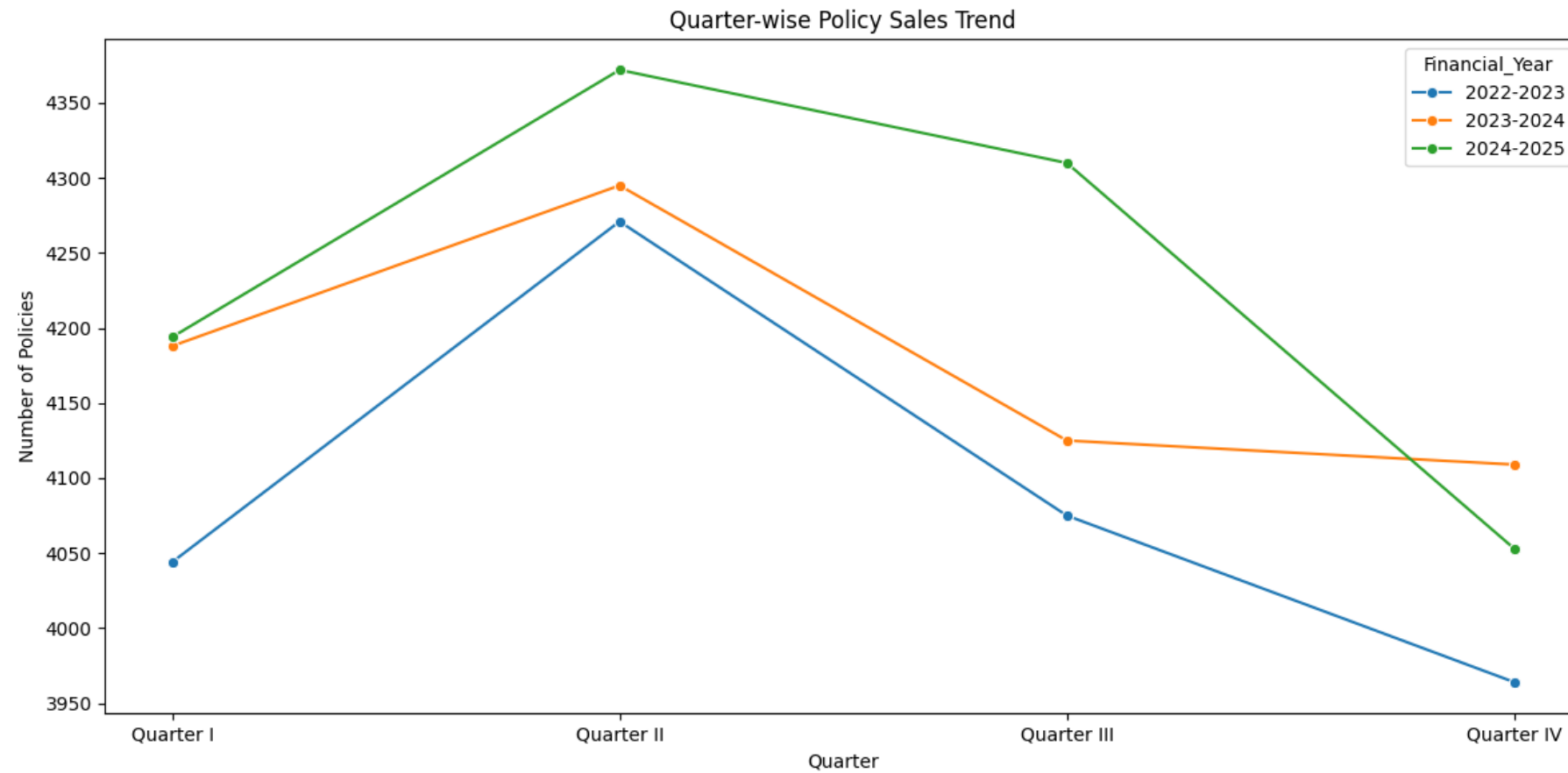
❖ In FY 2024-25 highest count of policies sold in July month and generated most of the business but sales dropped in month of February.

Total Policies sold over months in FY2024-25

Total Premium collected over months in FY2024-25
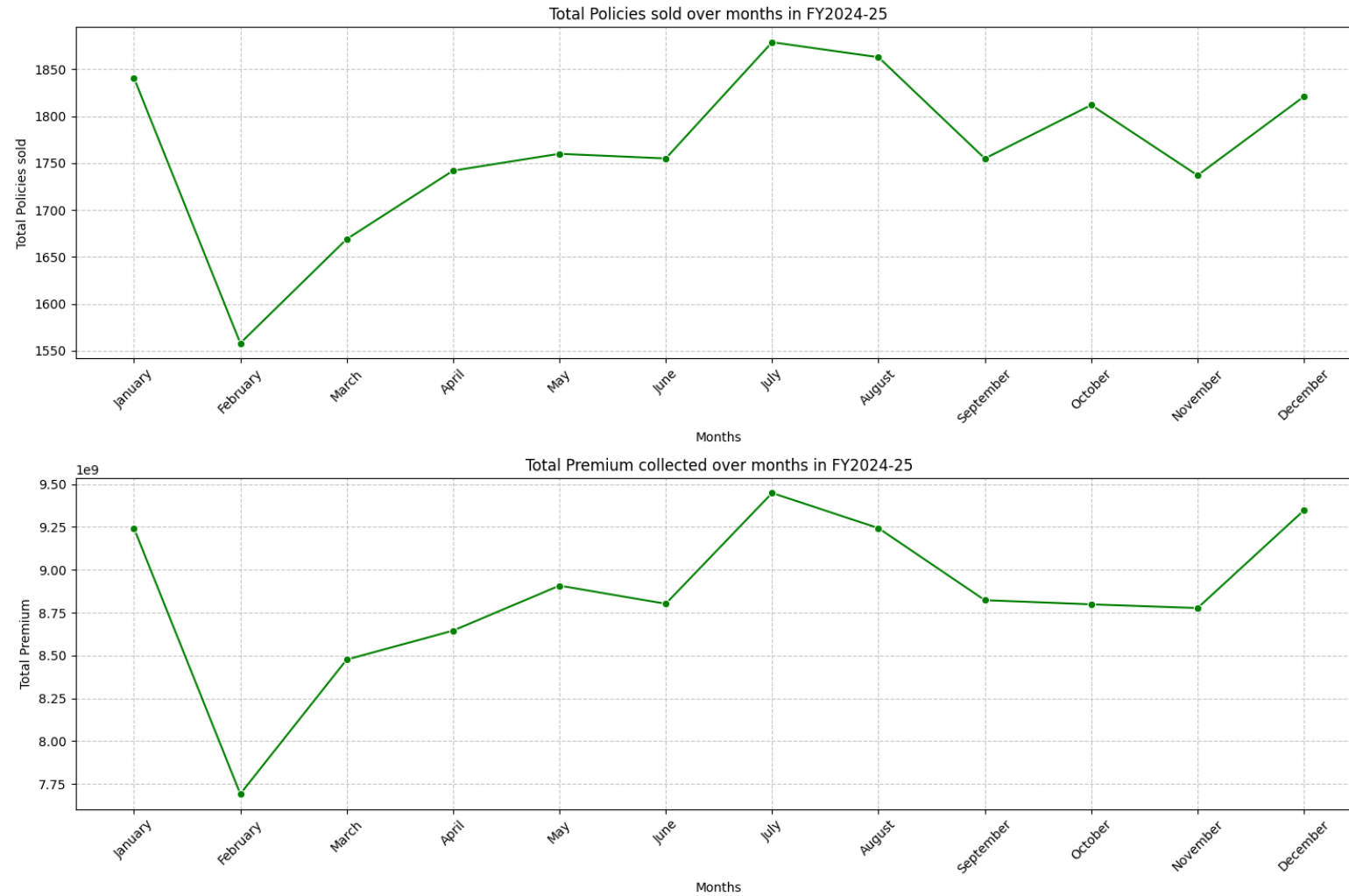
❖ In FY 2024-25 highest count of policies sold in August month but highest premium generated in July month, sales dropped in month of February.

Total Policies sold over months in FY2022-23

Total Premium collected over months in FY2022-23
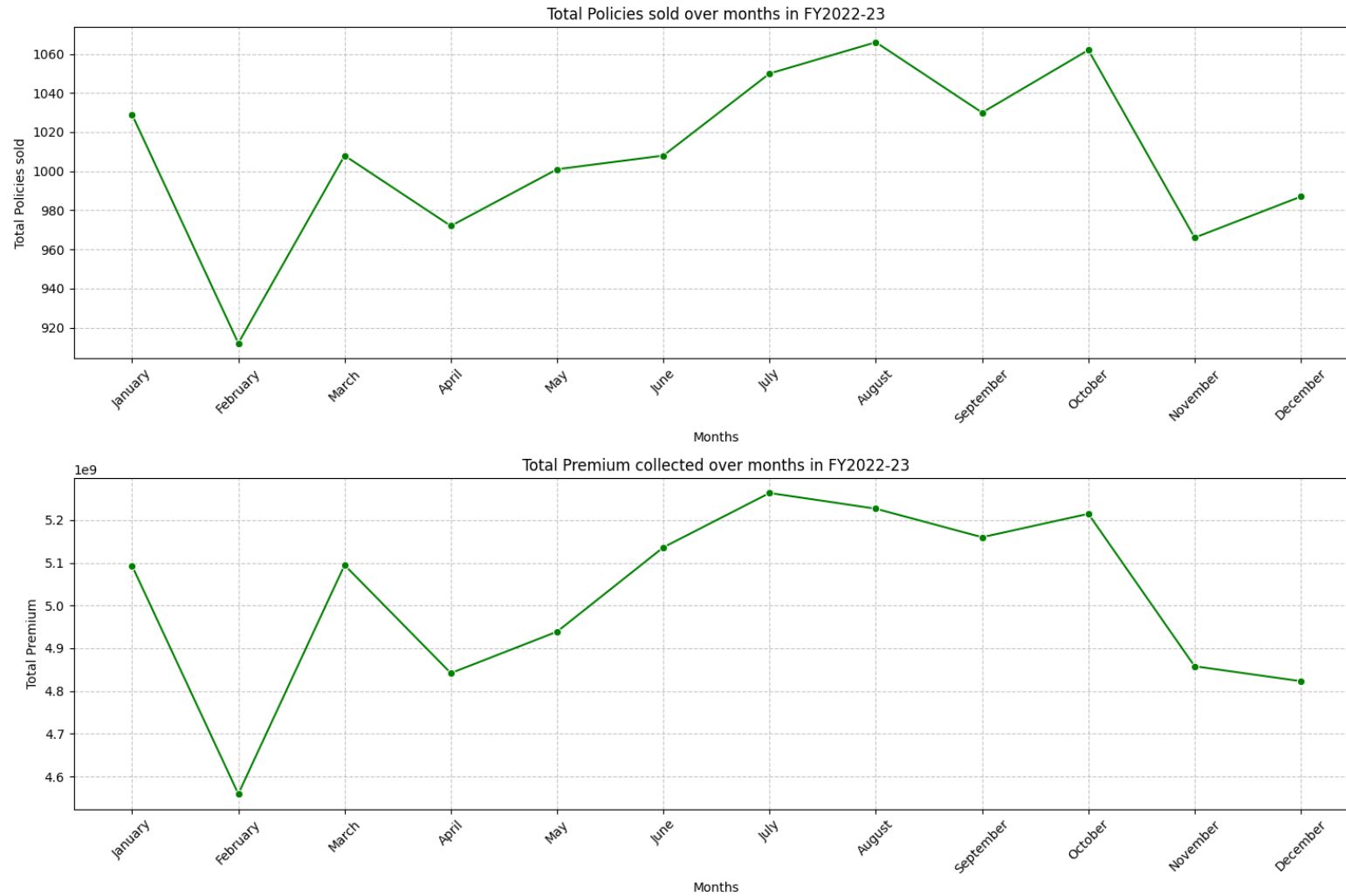
❖ In FY 2023-24 September month generated most of the business but sales dropped in month of November.
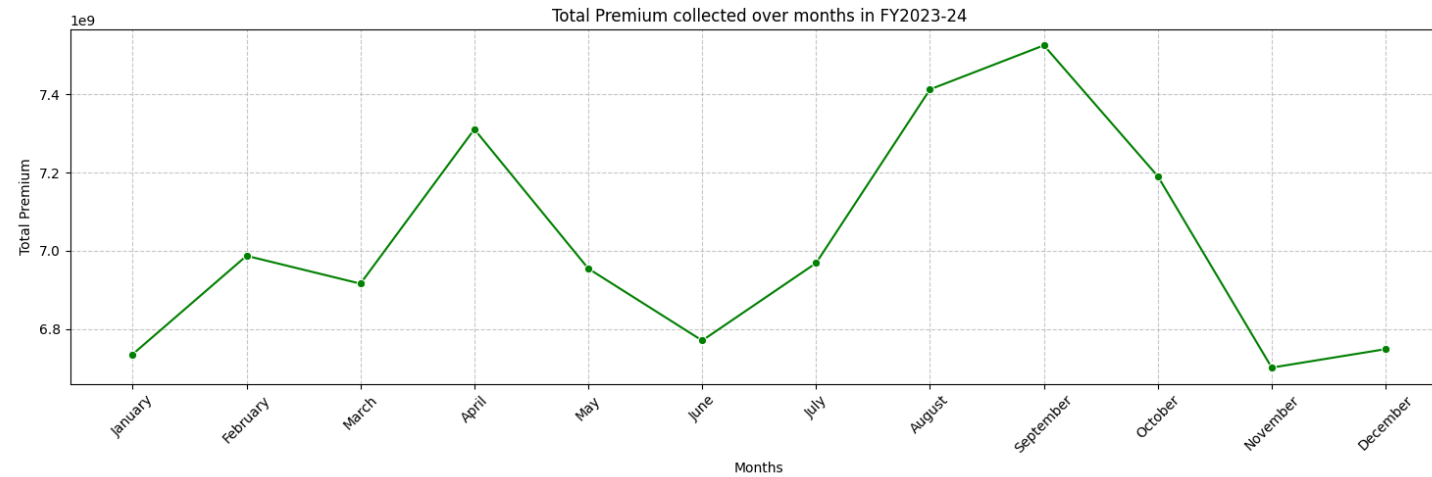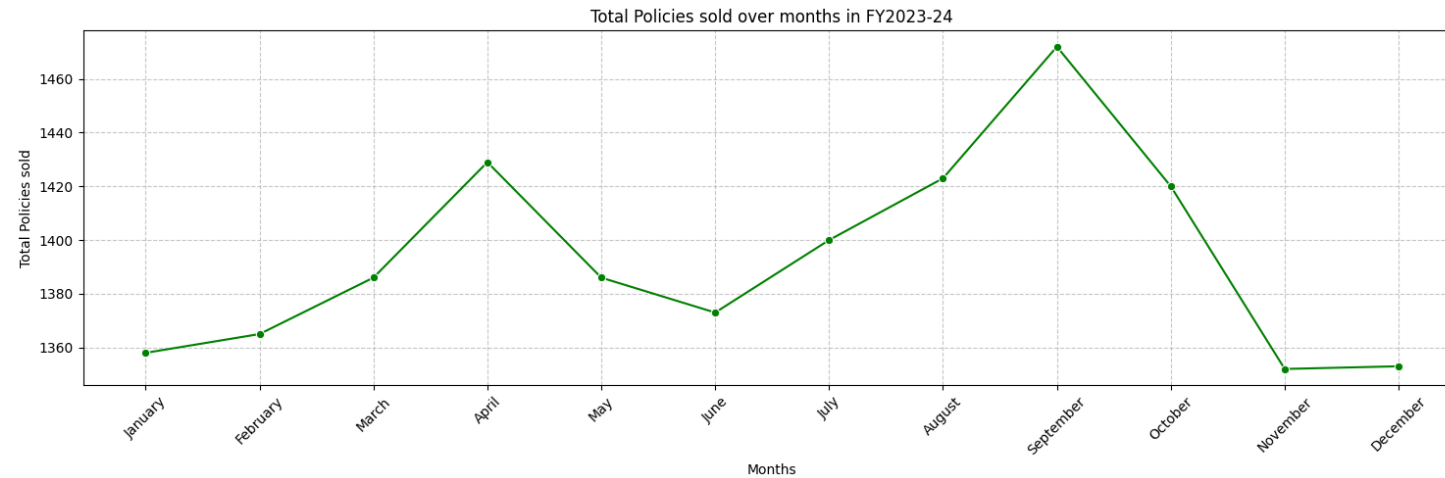
❖ In all the Quarters of Financial, Life policy type are sold most and Auto policy type is least

Quarterly Premium Contribution by Policy Type
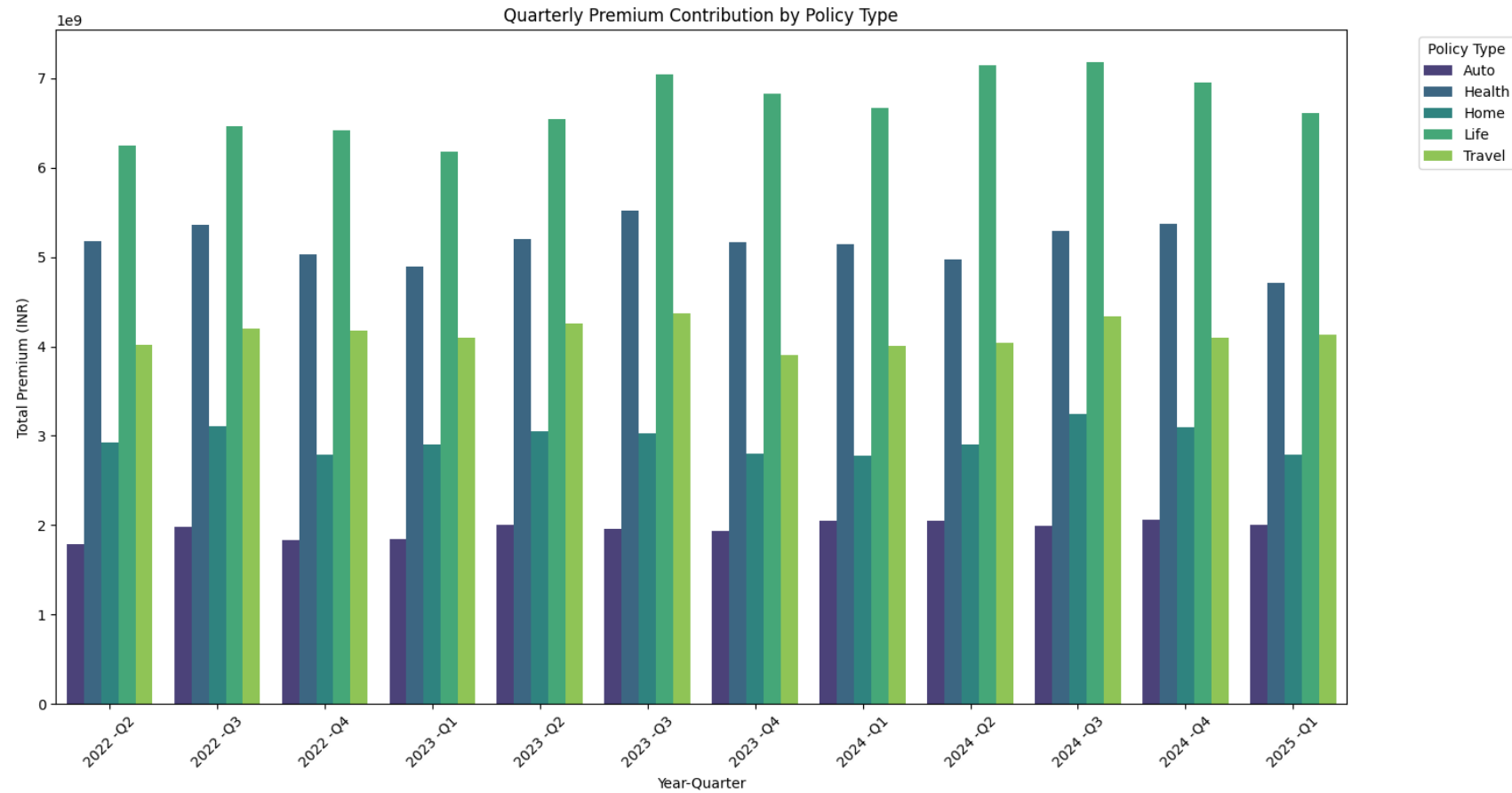
❖ The box plot shows that all policy types have a similar distribution of coverage amounts with medians around ₹1.1 million. Each category contains high-value outliers, indicating a few policies with significantly higher coverage.

Distribution of Coverage Amount by Policy Type
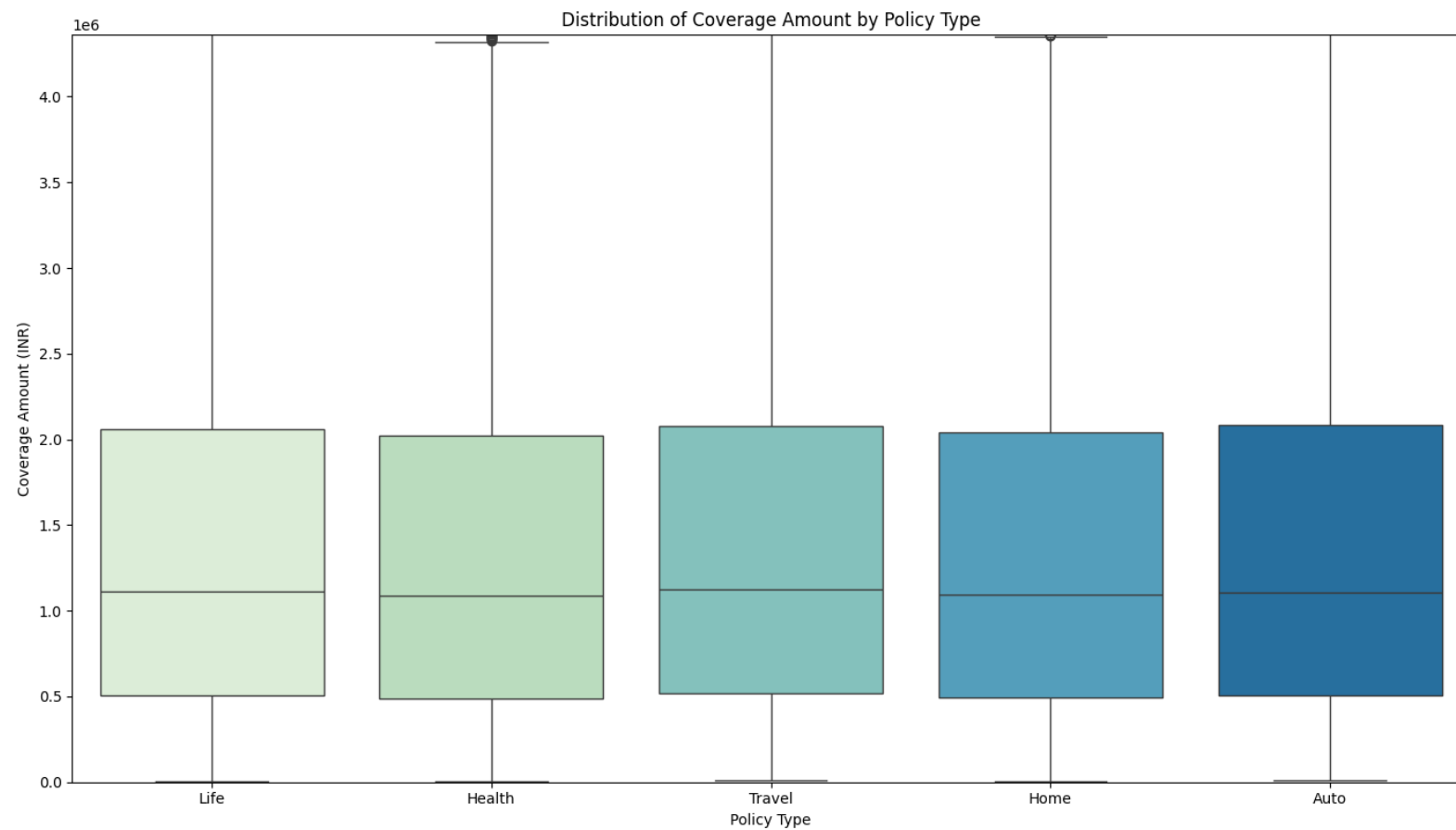
❖ The violin plot shows that Life insurance policies have the widest range and highest duration, with many lasting over 25 years. In contrast, Health, Travel, Home, and Auto policies are mostly concentrated between 1 to 3 years. Life insurance also has a bimodal distribution, indicating two common duration patterns short and long-term.

Distribution of Policy Duration by Policy Type
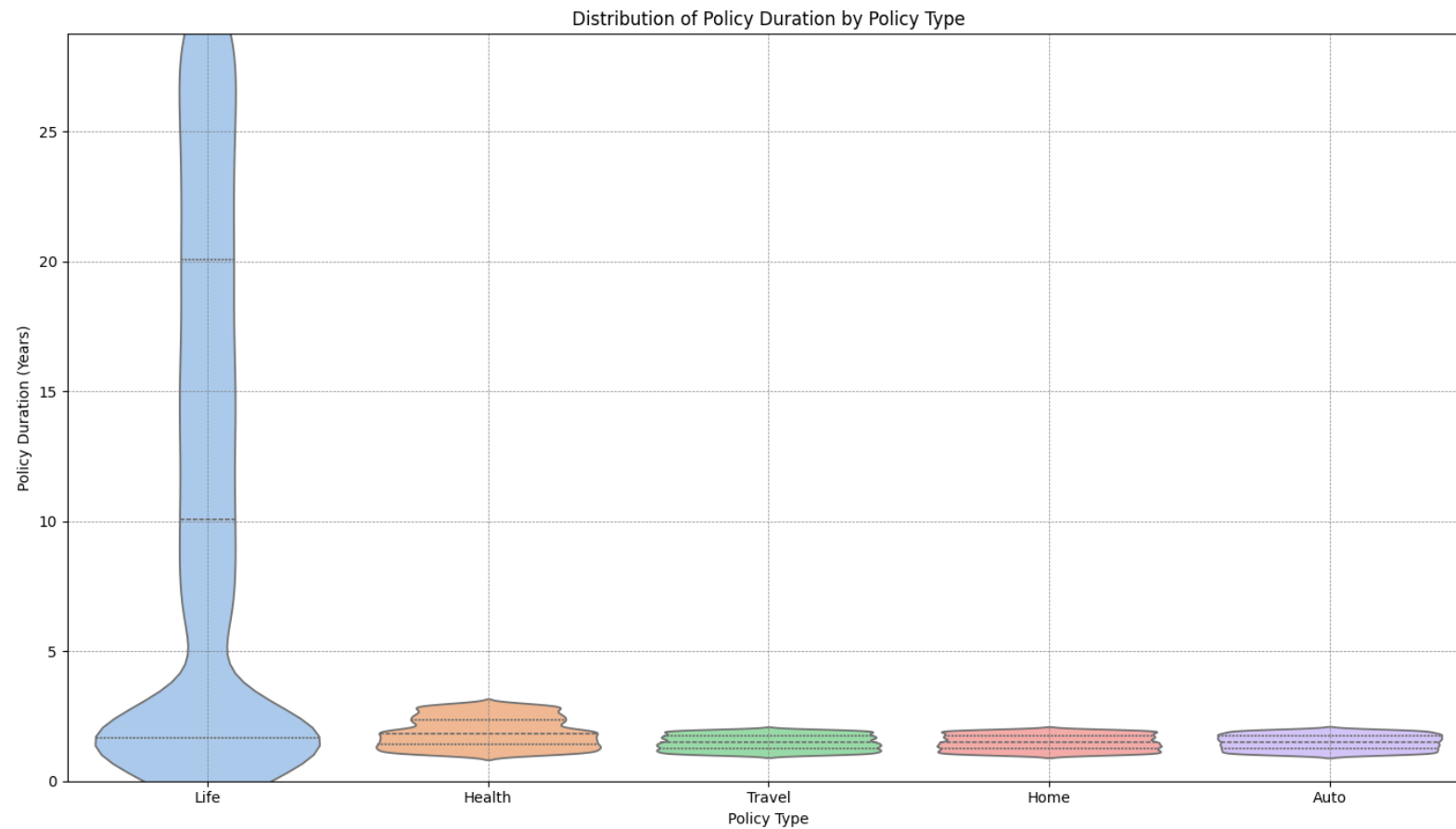
❖ Life insurance dominates across all sales channels, with the Agent channel showing the highest share (34%). Health policies are more balanced but also significant, especially through agents (28%). Auto policies are the least popular, especially in the Agent channel (7%). Travel policies are balanced in all the channels.

Distribution of Policy Types Across Sales Channels (Column Normalized)
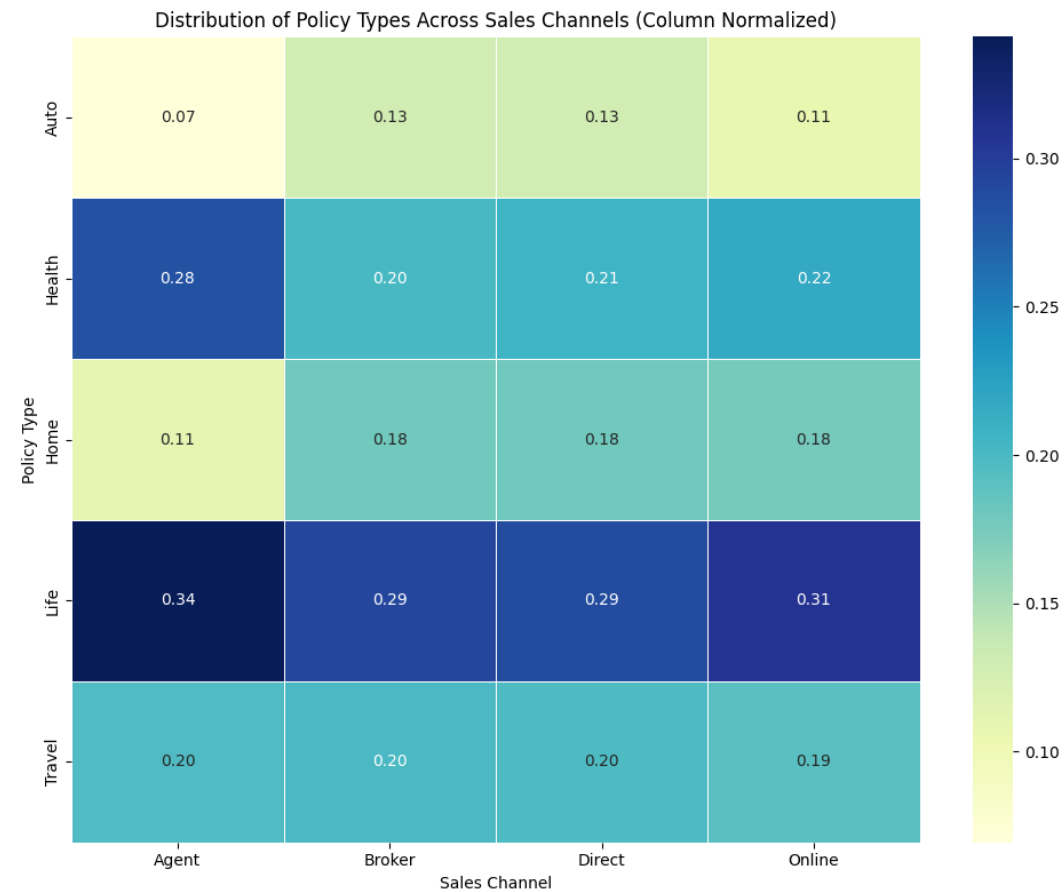
❖ All the genders are more focused towards life insurance and shares higher %age. Health policies are balanced in all the genders. Auto policies are least preferred by consumers.

Distribution of Policy Types Across Customer Gender (Column Normalized)

Policy Type	Female	Male	Other
Auto	0.09	0.09	0.10
Health	0.25	0.25	0.25
Home	0.14	0.14	0.14
Life	0.32	0.32	0.32
Travel	0.20	0.20	0.19
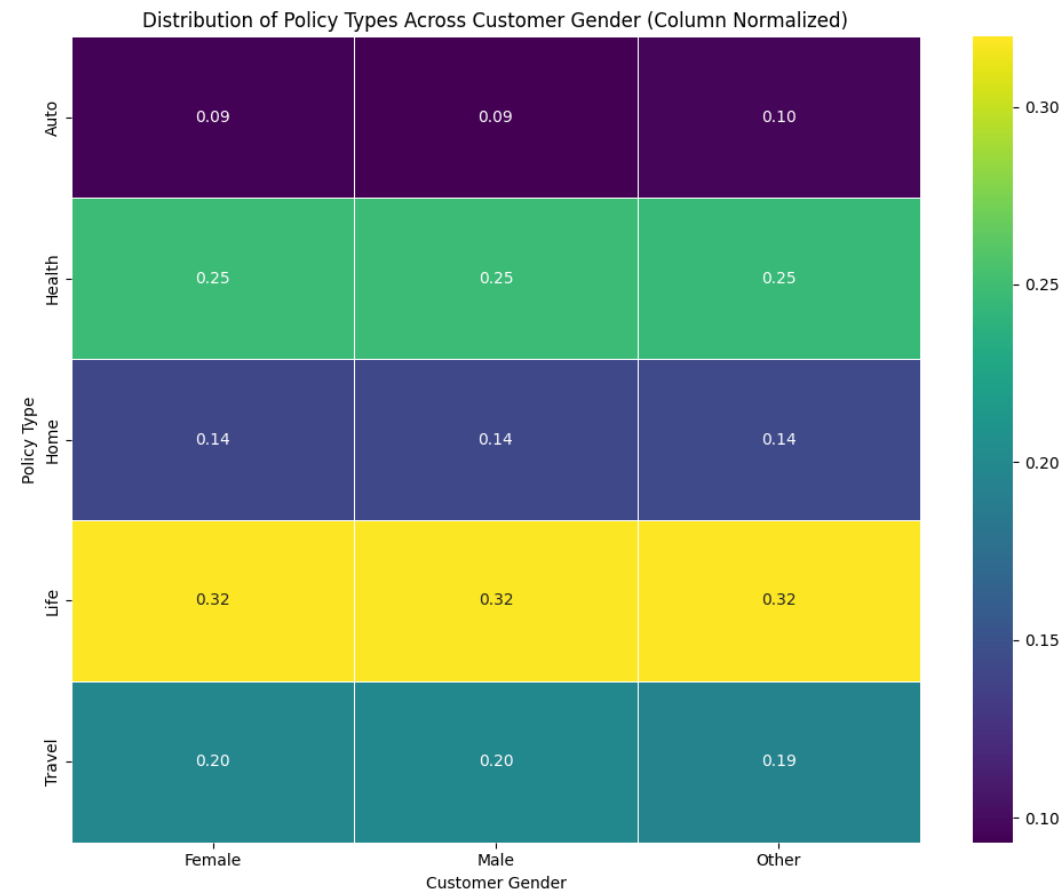
Customer Gender

❖ Agents, Broker, Direct channel focus in Male and Female whereas online channel focused to Male and Other genders.

Distribution of Sales Channel Across Customer Gender (Column Normalized)
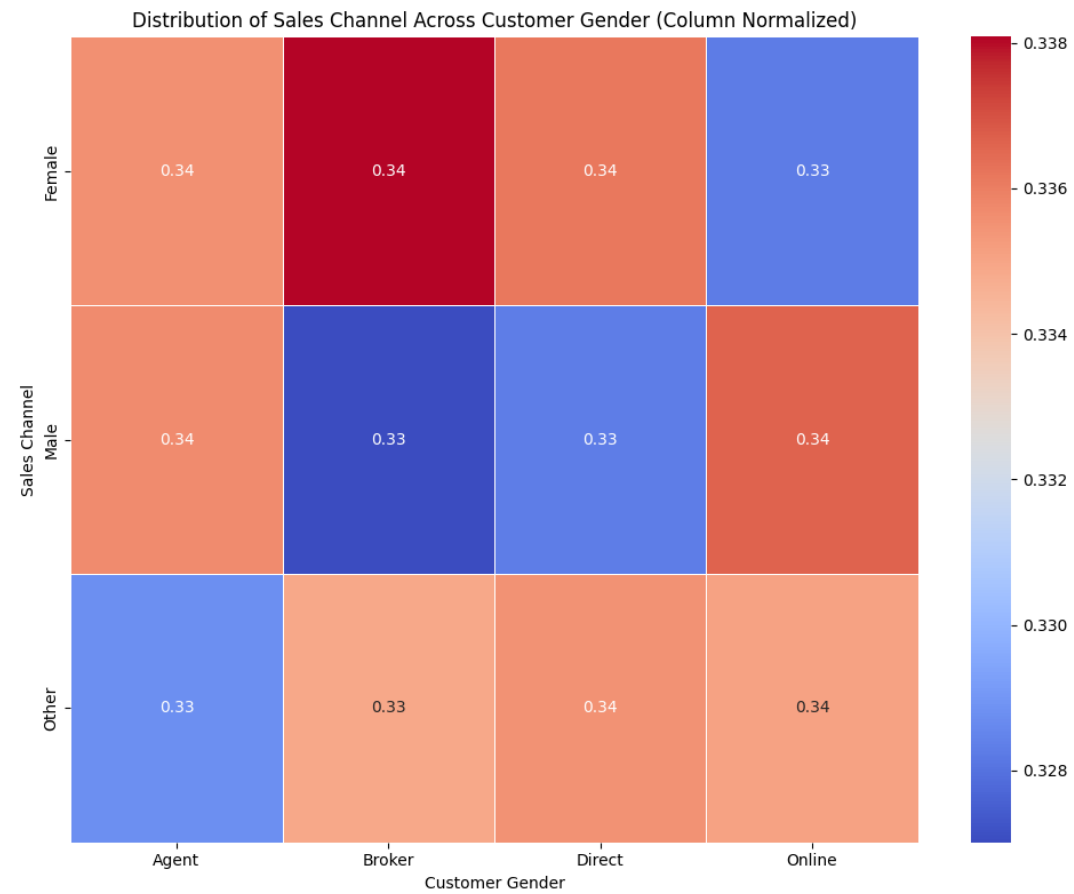
❖ This heatmap shows average claim probabilities across age groups and policy types.
Younger customers (18–24) have the highest claim rate for Auto policies (0.53), while seniors (65–74) lead in Home and Travel claims (both 0.53–0.54) and (18–24) least share in Home claims .
Most other values range between 0.48–0.51, showing moderate consistency across groups.

Average Claim Probability by Age Group and Policy Type