

Dataset Cleansing

This stage was focused on identifying the missing values and remove any duplicates. In order to achieve this Power Query was used and 9 duplicate lines were removed while missing values of Premium per Asset column were replace with the average value per policy. Then, missing values of Dwelling Limit column & Dwelling Deductible were replaced with zero. The reason for that, was that we wanted to have data consistency and these columns to be decimal numbers. Moreover, we were concluded to this option as there was average value which meant that there was a pattern for these policies.

When this was completed, Premium per Asset, Dwelling Limit & Dwelling Limit deductible were unpivoted so that we have all the rows as there were locations that were recorded in one row. In cases, where the postcode was different in these rows we created 3 different queries one per postcode and we appended them so that we get the final dataset.



Exploratory Analysis

For this stage we used the Value column filtered by Premium per Asset to get the distribution per state and carrier.

For the any patterns in terms of end and start date, we used column charts and specific trends were visualized.



Compliance Checks

On this stage, the most important things was to define the thresholds of Dwelling Limits and Dwelling Deductible. We made the assumption that not all of the policies have the same thresholds. Moreover, data such as the size of the property were missing which could give us better indication of the threshold.

However, the average of Dwelling limit and average Dwelling deductible were calculated plus and allowance which was randomly selected so that we don't split our sample in two.

Reccomedations of improvement can be found in Compliance Checks Analysis page.



Advance Analytics (1/2)

On this stage both excel and power bi were used.

Excel: The "data analysis" tool was used to create a linear regression between Premium per Asset and Dwelling Limit and Premium per Asset and Dwelling Deductible.

From the results, we understand that between Premium per Asset and Dwelling Deductible there is a linear regression. This is not strong as the Multiple R is equal with 0.27.

There is no linear regression between Premium Asset and Dwelling Limit.

Further information can be found in excel spreadsheet. Please note that the clean dataset was used after a series of transformation with aim of making the type of amount column into multiple column based on the distinct values of each row.



Advance Analytics (1/2)

Power Bi: Segmentation was conducted in Power BI for the policies based on the number of locations.

Firstly, Policy Type and Carrier were used to calculate the number of row which will give us how many policies per carrier. Then segmentation groups were created to show the distribution of our dataset.

Secondly, Policy Type and States were used to calculate the number of rows which will give us how many policies per carrier. Then segmentation groups were created to show the distribution of our dataset.

These were created in two different queries and then visuals on the Advanced Analytics page were created.

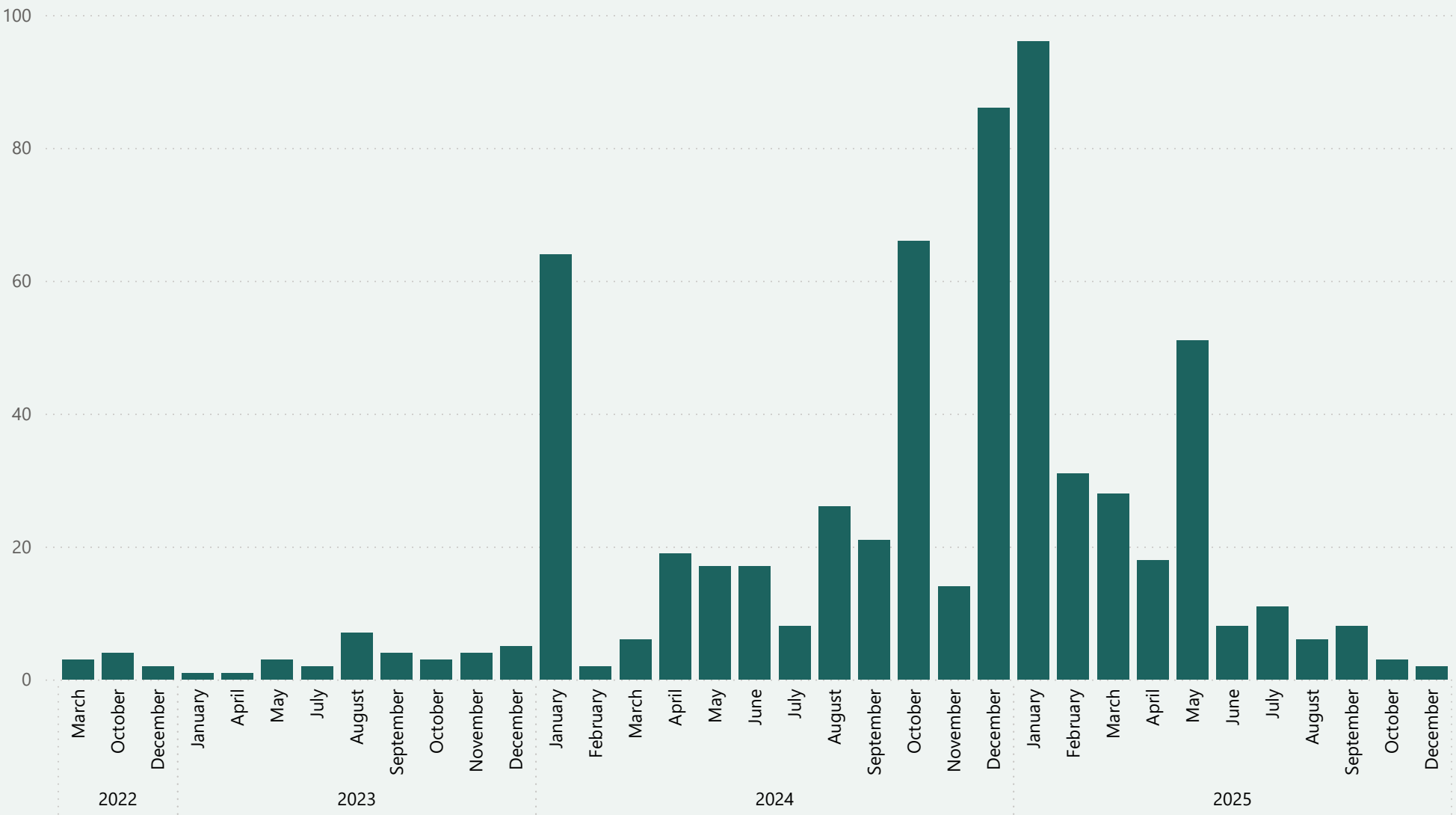


Deliverables

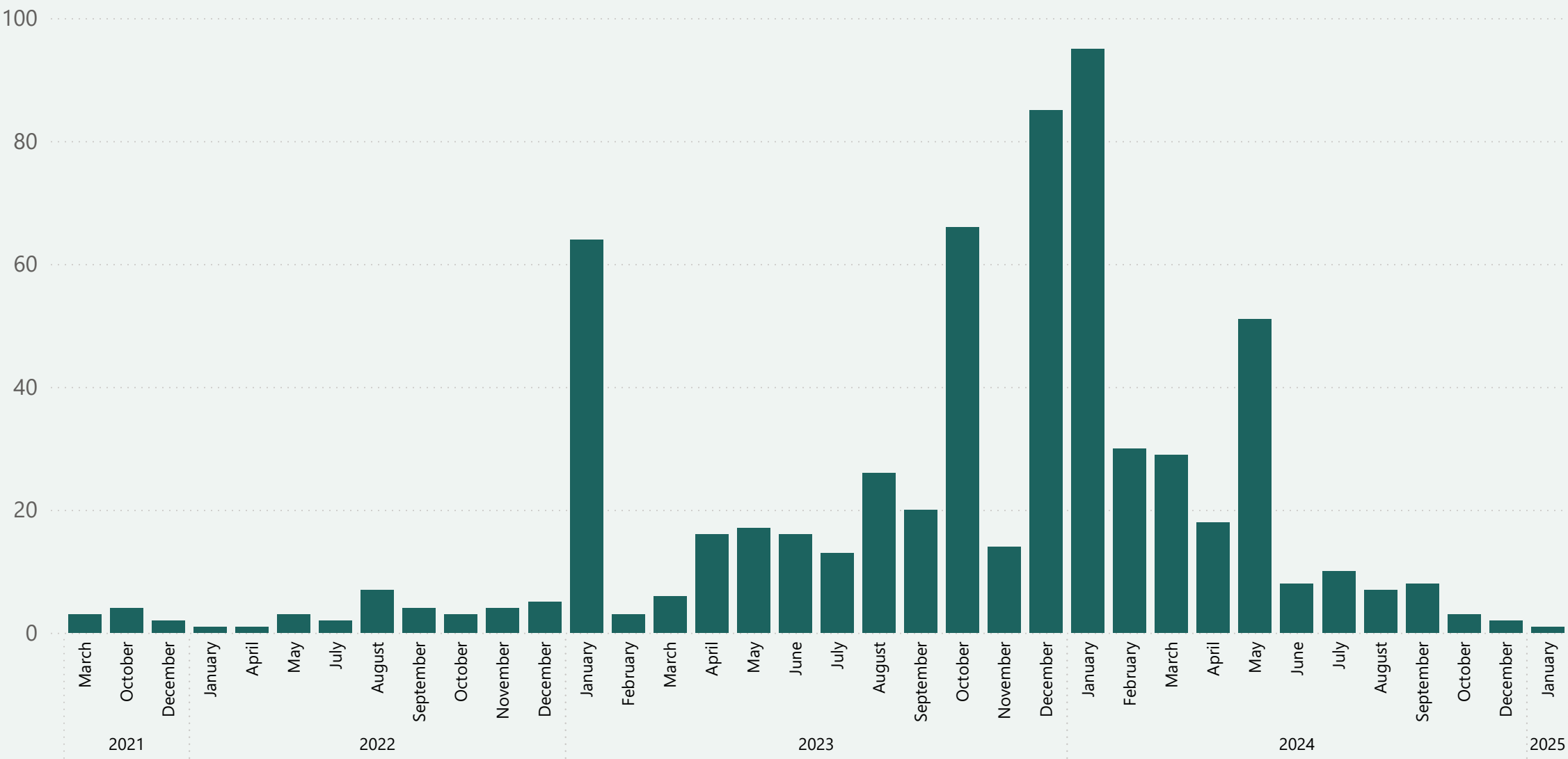
- 1) Power BI: pbi file with the queries and key visualisations
- 2) Pdf export: this is produced by the power bi
- 3) Excel: Here you can find the initial dataset along with the results of linear regression for each case
- 4) Clean Dataset: here you can find the clean dataset along with the transformations required so that it is ready to be used for the linear regression modelling.



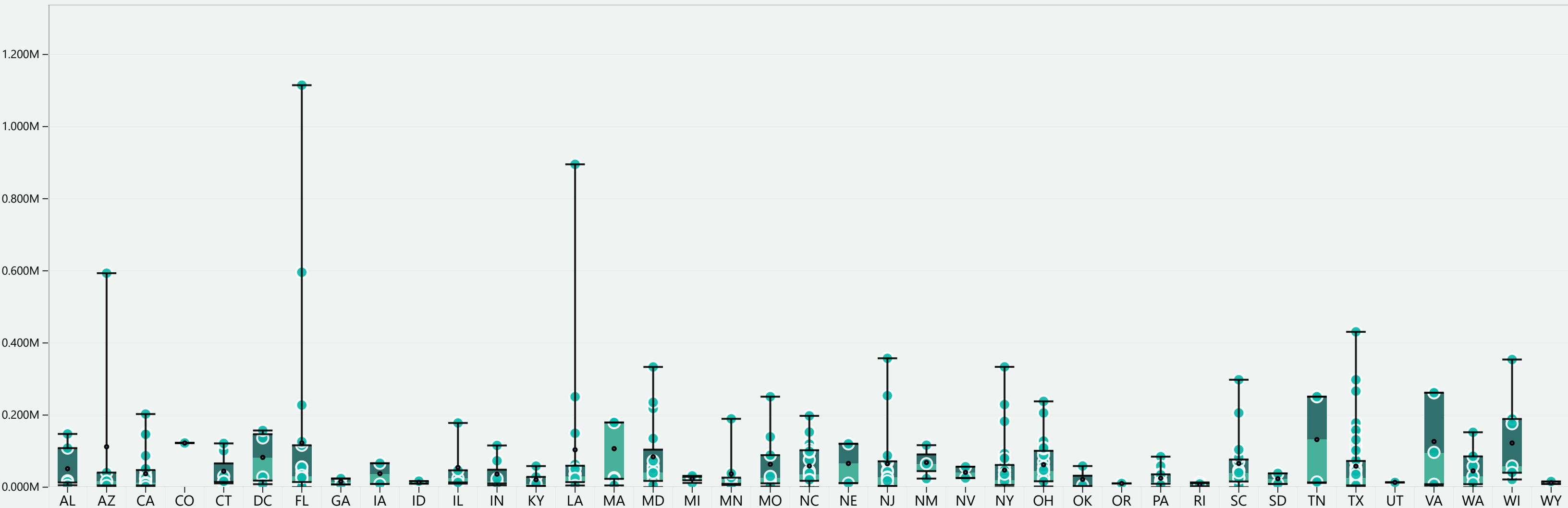
No. Policies Ending Each Month & Year



No. Policies Starting Each Month & Year



Distribution Premium per Asset by State



Commentary

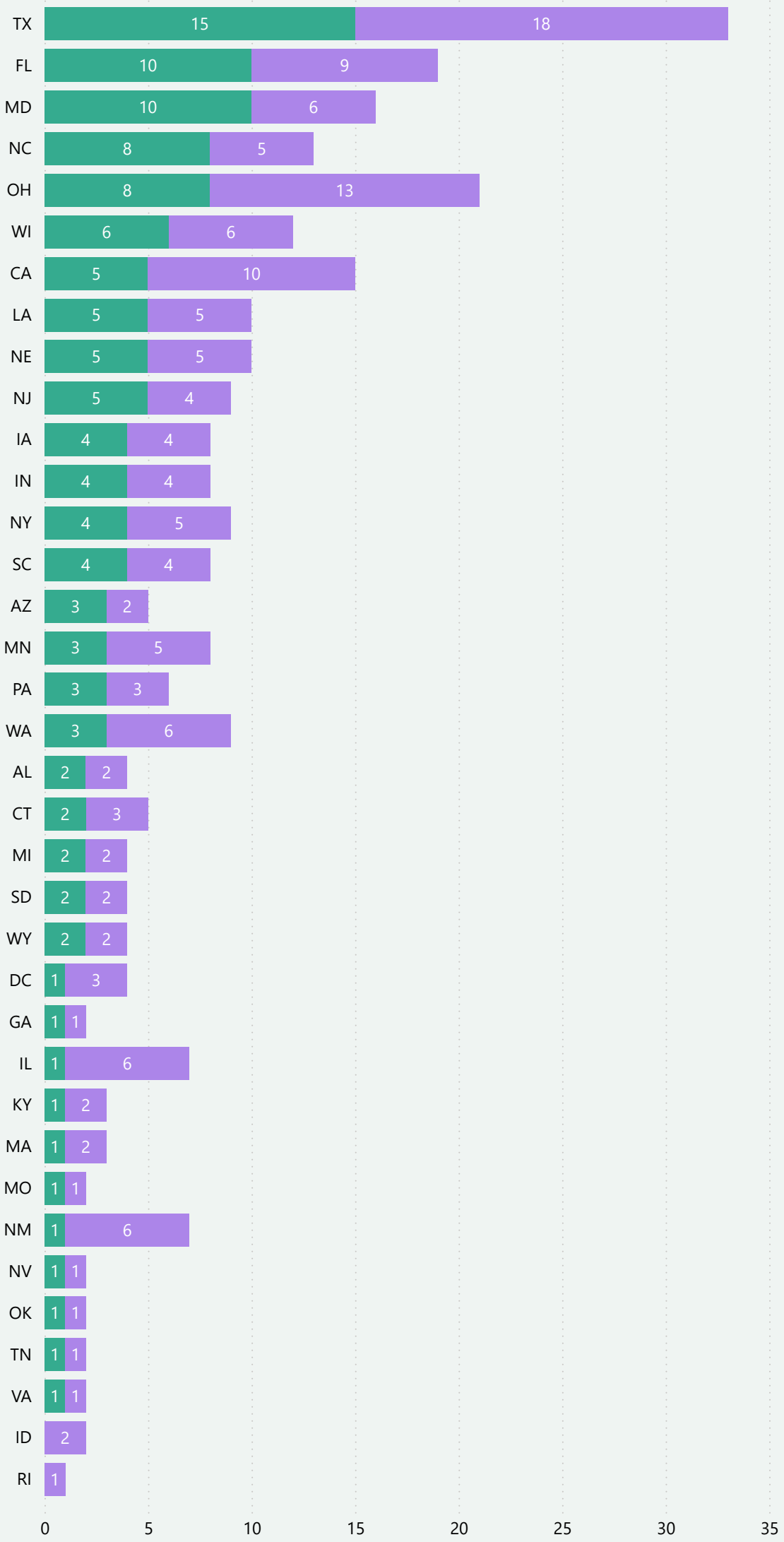
The boxplot shows the distribution of Premium per Asset across different states. There are variations in premium amounts between states, with some states exhibiting higher premiums and larger spreads.

The bar plots for policy begin and end dates indicate the frequency of policies starting and ending each month. There are peaks in certain months, which may suggest seasonal trends or business cycles in policy insurance and termination.



Non Compliant Distribution per State

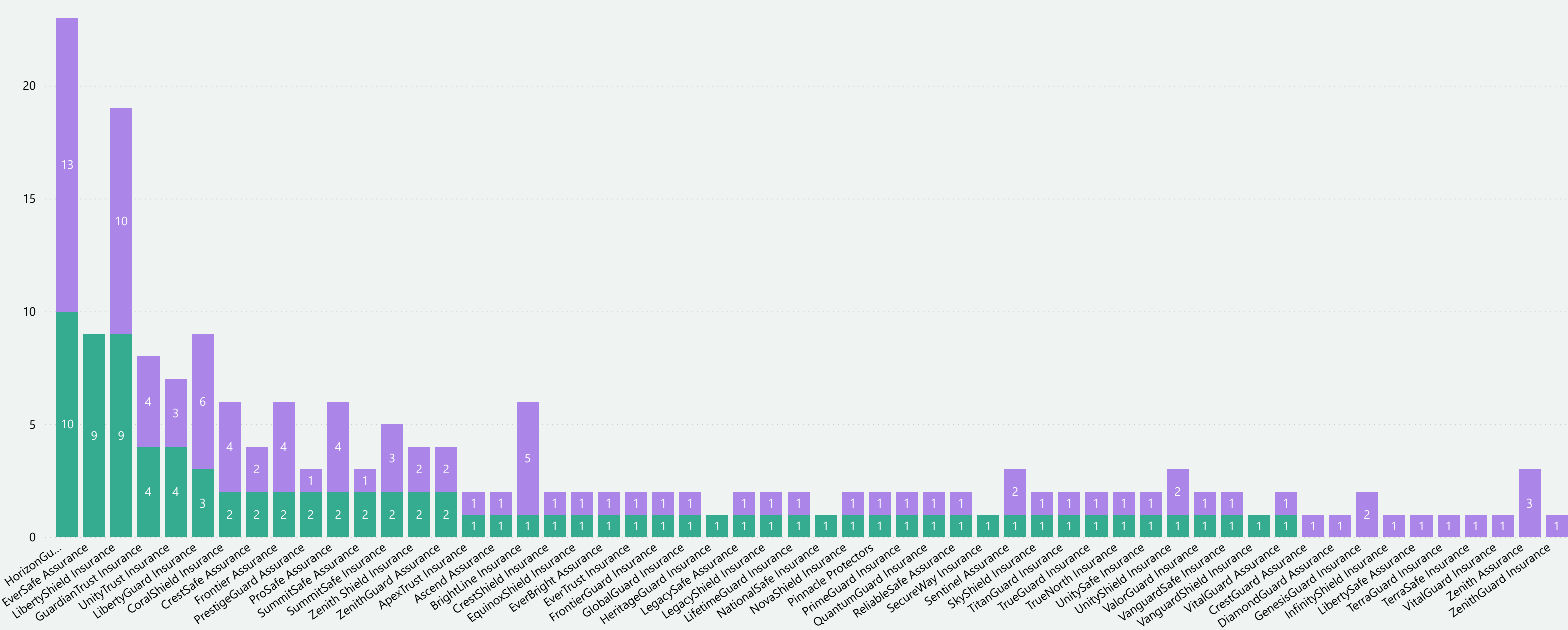
● Non Compliant - Dwelling Limit ● Non Compliant - Dwelling Deductible



Product Type	Non Compliant-Dwelling Deductible	Compliant-Dwelling Deductible	%	Non Compliant-Dwelling Limit	Compliant-Dwelling Limit	%
Auto Liability Policy		48	0.0%		48	0.0%
Boiler and Machinery / Equipment Breakdown Policy		9	0.0%		9	0.0%
Business Owners Package (BOP) Policy	6	39	13.3%	3	42	6.7%
Commercial Property Policy	30	294	9.3%	30	294	9.3%
Earthquake Policy		6	0.0%		6	0.0%
Flood Policy		30	0.0%		30	0.0%
General Commercial Package Policy	36	440	7.6%	50	426	10.5%
General Liability Policy		308	0.0%		308	0.0%
Terrorism Liability Policy		15	0.0%		15	0.0%
Terrorism Package Policy	1	8	11.1%	2	7	22.2%
Terrorism Property Policy		90	0.0%	12	78	13.3%
Umbrella Liability Policy		563	0.0%		563	0.0%
Wind/Hail Policy		9	0.0%	1	8	11.1%

Non Compliant Distribution per Carrier

● Non Compliant - Dwelling Limit ● Non Compliant - Dwelling Deductible



Commentary

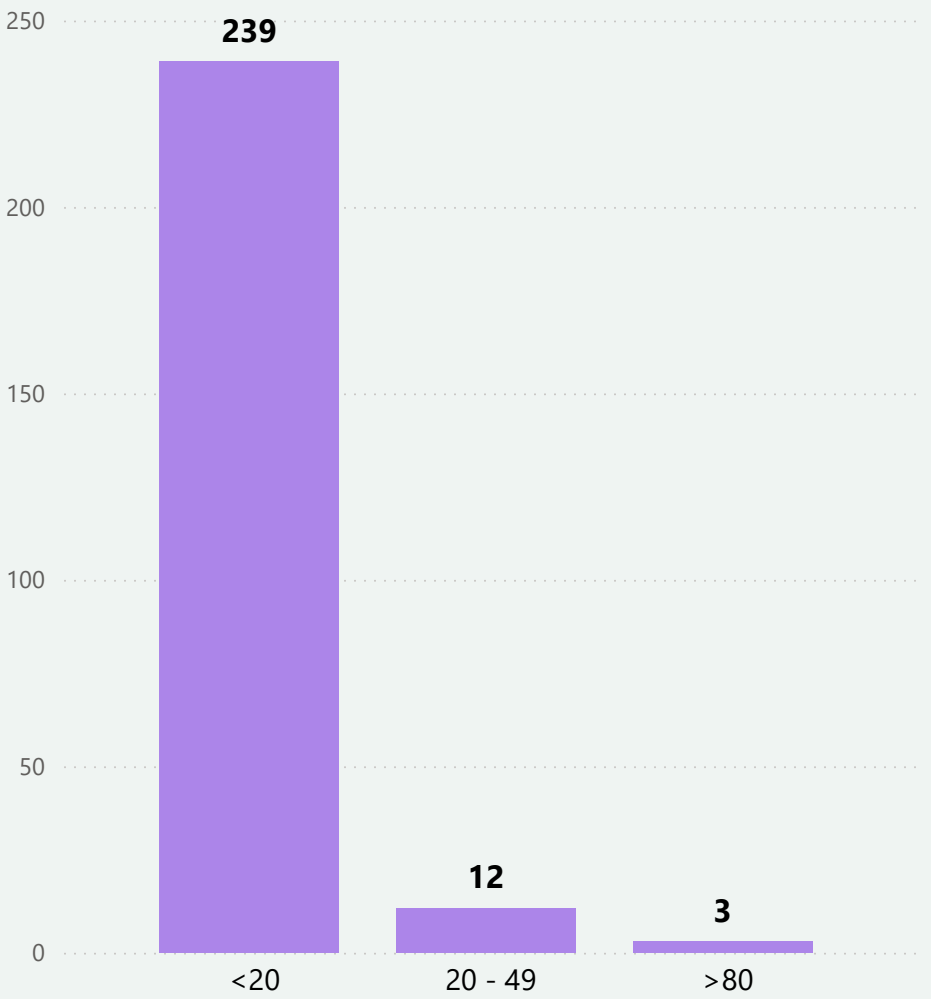
Assumptions:  
1) Threshold for Dwelling Limit was set as the average of Dwelling Limit amount in the dataset plus **\$1,500,000**  
2) Threshold for Dwelling Deductible t was set as the average of Dwelling Limit Deductible amount in the dataset plus **\$1,000**  
3) For the policies with 0 dwelling limits, an assumption was taken that they do not require a limit therefore not threshold was defined.  
Details:  
1) The charts in this page aims to show the distribution of non-compliant policies per state and per carrier  
2) The table below shows the compliant and non compliant policies and compares them. An assumption was taken that if the non-compliant policies is above 10% of total policies then further investigation is needed

Recommendations:

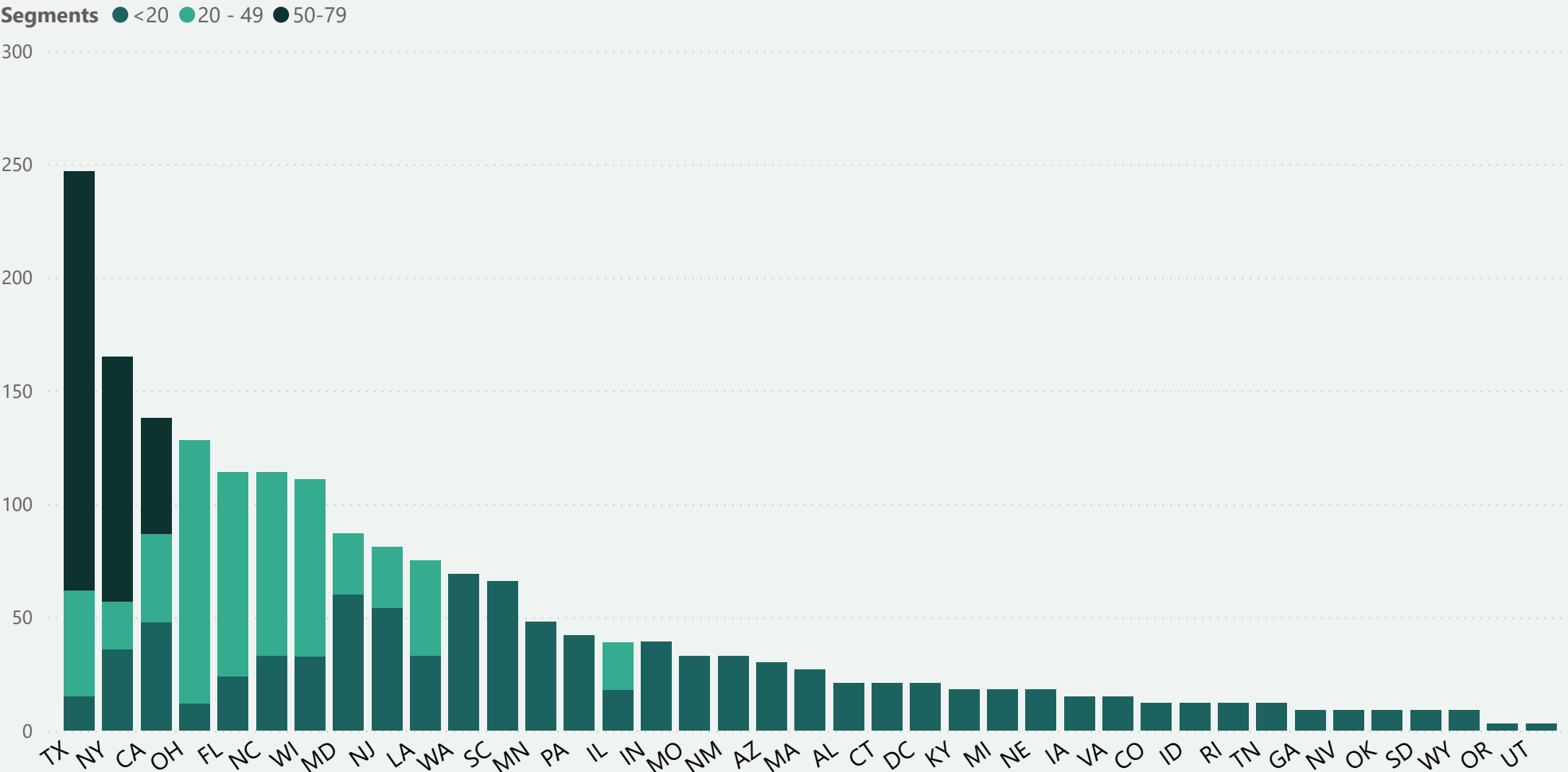
While non compliant policies are not a majority, there are cases like Terrorism Liability Policy where deductible and dwelling limit are not compliant in >10% of the cases. It also looks like that TX and FL have the most cases of non compliance. This could be happening due to several reasons but briefly let's mention how we can improve on this:  
1) turn our attention on how the data are populated and seek opportunities for improvement  
  
2) Use of reports and dashboard that will visualise non-compliance cases in a regular basis and allow people to update the data or investigate the reasons why we have non-compliance  
  
3) integrate more data so that we achieve better accuracy. For example, size of the property. This will give us better indication of how much the dwelling limit and dwelling ded limit is supposed to be. On top of that specific threshold will need to be implemented that those should be known.



No Policies per Carrier in Segments



No Policies per State in Segments



Commentary

This page contains 3 visuals that represent the policy segmentation by type, and carrier.

In order to achieve this, the dataset was grouped by Carrier and Policy type and then segments were created. These were based on the count of policies per carrier and type. These segments are the following:

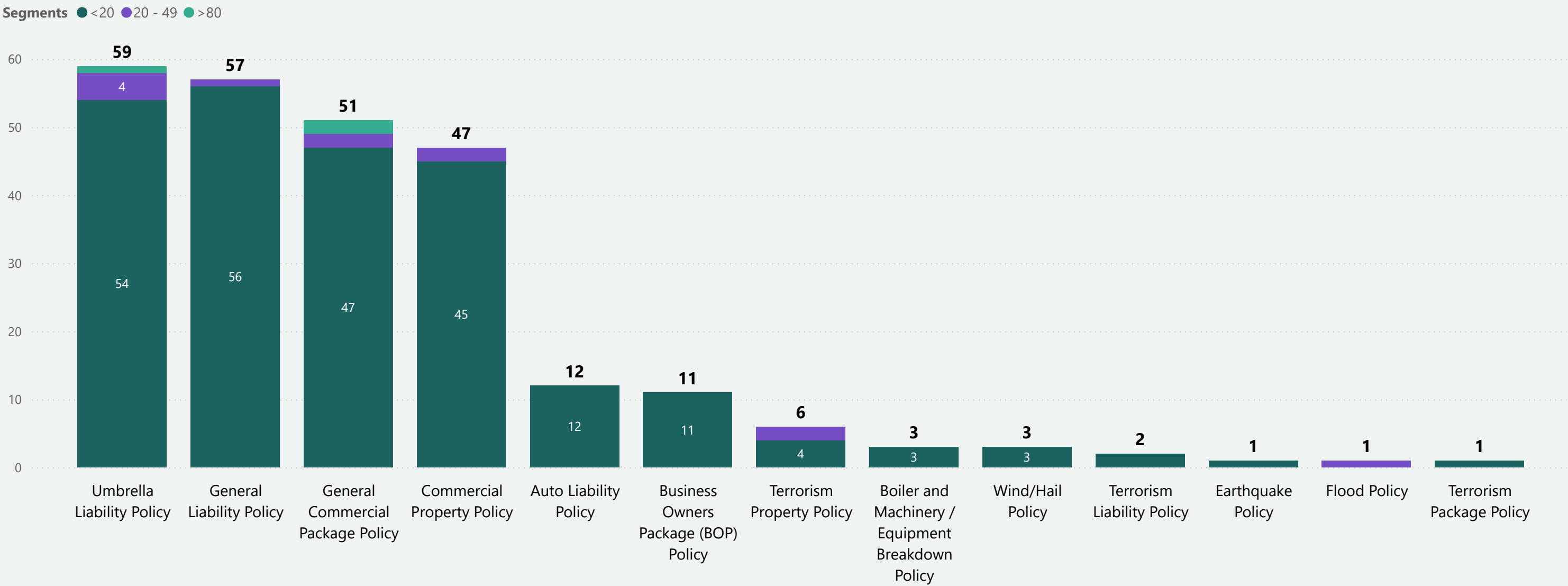
- 1) <20
- 2) 20-49
- 3) 50-79
- 4) >80

The majority of policies belong to the first segment <20 while the Umbrella Liability Policy, General Commercial Package Policy, Commercial Property Policy and General Liability Policy are the most popular within the dataset.

In terms of the carriers, a list of the top 20 carriers are presented with BrightShield Insurance to be the most popular along with VanguardShielded Insurance & HorizonGuard Insurance to be very close in number of policies.

The same segments were created using States & policy type where TX has the most locations with 185 policies active in 50-79 locations.

No Policies per Type & Carrier in Segments



Top 20 Carriers by No. Policies

Carrier	No. Policies
BrightShield Insurance	120
VanguardShield Insurance	117
HorizonGuard Insurance	114
TrueSafe Insurance	105
EverBright Assurance	48
GuardianTrust Insurance	45
StarSafe Assurance	42
UnityTrust Insurance	42
BrightLine Insurance	39
Frontier Assurance	39
LibertyShield Insurance	39
EquitableGuard Insurance	36
EverSafe Assurance	36
LifetimeGuard Assurance	36
Pinnacle Protectors	33
LibertyGuard Insurance	30
AtlasGuard Insurance	27
CoralShield Insurance	26
VictoryGuard Insurance	24
GlobalGuard Insurance	21