**Jelil Adebayo Adetoro**

**Task 1:**

There is no firm that meets all the requirements. However, if we break it down, we discover the following:

Top ten firms based on GWP:

Firms

Firm 210 265786.579808

Firm 4 214185.524094

Firm 311 85551.506309

Firm 247 77714.678591

Firm 105 66653.808223

Firm 26 52013.206571

Firm 34 51192.871255

Firm 73 47211.048347

Firm 7 43080.871198

Firm 151 39270.965362

A graph of blue bars

Description automatically generated

Top ten firms based on NWP:

Firms

Firm 210 177057.113250

Firm 4 145559.914004

Firm 26 112045.243873

Firm 105 83941.570969

Firm 311 42564.748795

Firm 247 41327.649473

Firm 199 38089.815171

Firm 151 34843.746508

Firm 7 31888.489669

Firm 73 31638.362726

A graph of blue bars

Description automatically generated with medium confidence

Top ten firms based on SCR Coverage Ratio:

Firms

Firm 216 9.993029e+08

Firm 131 9.269834e+08

Firm 1 5.588700e+07

Firm 66 7.162000e+06

Firm 320 6.735564e+06

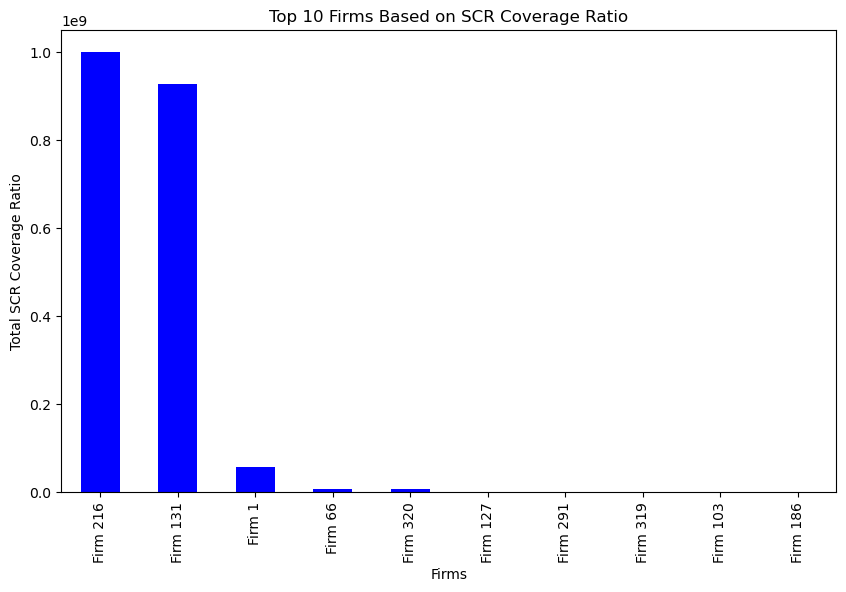
Firm 127 5.685685e+05

Firm 291 4.663403e+02

Firm 319 1.544702e+02

Firm 103 1.350224e+02

Firm 186 1.259393e+02



Top ten firms based on Gross claims incurred:

Firms

Firm 216 16597.874490

Firm 112 14735.128670

Firm 105 14090.246087

Firm 283 9956.683824

Firm 234 9244.238864

Firm 52 9132.591276

Firm 286 8205.734159

Firm 17 7599.959587

Firm 22 6754.931991

Firm 280 6137.748500

A graph of a number of firm

Description automatically generated with medium confidence

A graph of different colored lines

Description automatically generated

Top ten firms based on Net combined ratio:

Firms

Firm 99 -73450.271014

Firm 55 -224.127946

Firm 50 -34.158725

Firm 146 -18.018484

Firm 191 -15.805683

Firm 251 -6.258193

Firm 151 -3.745706

Firm 49 -3.219590

Firm 308 -3.141834

Firm 262 -0.064373

A graph with numbers and a number of firm names

Description automatically generated with medium confidence

Top ten firms based on total assets:

Firms

Firm 311 1.494586e+06

Firm 210 1.010947e+06

Firm 10 1.010333e+06

Firm 105 8.238229e+05

Firm 34 7.256552e+05

Firm 247 4.996297e+05

Firm 26 4.267759e+05

Firm 101 4.049999e+05

Firm 4 3.976302e+05

Firm 199 3.915218e+05

A graph of green bars

Description automatically generated

Also, looking at the common firms that have been among the top 100 of most of the metrics, we noticed the following firms:

* Firm 89, 25, 216 and 131

**Task 2:**

I decided to use unsupervised learning techniques like clustering to find groups of similar firms in the data. This allows identifying firms that behave in an outlier manner compared to most others.

A chart with numbers and dots

Description automatically generated

Firms further from their cluster centroid are more of the outliers. The biggest outlier firms are:

* Firms 4, 247, 210, 105, 26, 101, 151, 311, 7 and 10 etc.

**Task 3:**

We can utilize several Microsoft Azure cloud services in building an end-to-end pipeline for data processing and analytics. The following is a high-level outline of an Azure service-based batch processing pipeline:

**Data Ingestion:** We can use Azure Data Factory to import data from several sources such as CSV, RDBMS, JSON etc., including those on-premises and in the cloud.

**Data Storage and Archiving:** We can use Azure Blob Storage to save data files such as CSV and JSON for further processing. Azure Blob Storage gives the freedom to store massive volumes of unstructured data at a lower cost.

**Processing Data:** Azure Databricks is an analytics platform that is optimized for Azure and is built on Apache Spark. We can use it to execute analytics at scale, prepare data for analysis, and create features. Azure Databricks is compatible with SQL, R, Python, and Scala.

**Transforming Data:** We can also use Databricks' notebooks to clean up data, apply transformations, and get it ready for analysis.

**Machine Learning and Data Analysis:** We can utilize Azure Machine Learning for advanced analytics and machine learning. Azure Machine Learning Studio provides an environment where we can train, develop, and deploy models and deploy inference pipelines for use.

**Information Visualization:** We can make use of Power BI to create and deploy interactive dashboards and reports after the data has been processed and analyzed. It can offer real-time information by connecting directly to the above Azure services.

**Organization of Data and Processes:** Azure Data Factory can be utilized for both data intake and the orchestration of complicated operations. Azure Data Factory allows us to plan and execute a data processing pipeline.

**Monitoring and Logging:** We can monitor the status and performance of the whole process flow with Azure Monitor and Azure Log Analytics. Azure Monitor enables us to know how the Azure resources are running while we use Log Analytics to query and analyze logs.

**Security and Compliance:** With Azure Key Vault, we can protect sensitive information utilized by cloud services and apps, including cryptographic keys. With Azure Policy, we can define and enforce policies to ensure compliance with organizational standards.

**Other Considerations.**

We can also make use of other Azure tools and services if necessary. These can include Azure Functions, Azure SQL Database, Azure Synapse Analytics, Azure Data Lake Storage, Azure Firewall, Azure Virtual Network etc.