**Data Sources**

* **Data.gov**: Provides retail sales data, economic indicators, and consumer spending statistics.
* **IPUMS**: Offers demographic and socioeconomic data.
* **Data.Seattle.gov**: Supplies local retail performance data, including foot traffic and sales figures.

**Data Ingestion**

* **ETL Processes**: Extract, Transform, Load (ETL) processes are used to collect data from the various sources. This involves:
  + **Extraction**: Pulling data from Data.gov, IPUMS, and Data.Seattle.gov using APIs and data feeds.
  + **Transformation**: Cleaning and standardizing the data formats, handling missing values, and removing duplicates.
  + **Loading**: Importing the cleaned data into the cloud database.

**Data Storage**

* **Cloud Database**: A secure cloud-based database is used to store the ingested data. This database is accessible to all team members and supports regular backups to ensure data integrity and security.

**Data Processing**

* **Data Cleaning**: Further cleaning of the data to ensure high quality, including normalization and standardization.
* **Data Joining**: Merging datasets on common keys such as location, time, and demographic attributes to create a unified dataset.
* **Data Transformation**: Creating new calculated fields, aggregating data as needed for analysis, and preparing the data for analytical processes.

**Data Analysis**

* **Statistical Analysis**: Applying regression analysis to identify trends and correlations between e-commerce growth and brick-and-mortar store performance.
* **Geospatial Mapping**: Using geospatial analysis to map e-commerce growth against brick-and-mortar sales, providing visual insights into the data.
* **Other Techniques**: Employing additional statistical techniques to quantify the impact of e-commerce on different retail segments.

**Data Output**

* **Reports**: Generating detailed reports that summarize the findings and provide insights into the impact of e-commerce on brick-and-mortar stores.
* **Visualizations**: Creating visualizations such as charts, graphs, and maps to illustrate the data and findings.
* **Dashboards**: Developing interactive dashboards that allow stakeholders to explore the data and insights in a user-friendly format.
* **Strategic Recommendations**: Formulating strategic recommendations for retailers based on the analysis, helping them adapt to the changing retail environment.

A diagram of a data flow

Description automatically generated