Real Estate Sales & Market Analysis

A snapshot of sales, property trends, and client behavior.It helps in understanding key metrics and offers historical trends and forecasts for sales, along with breakdowns by location and property type.

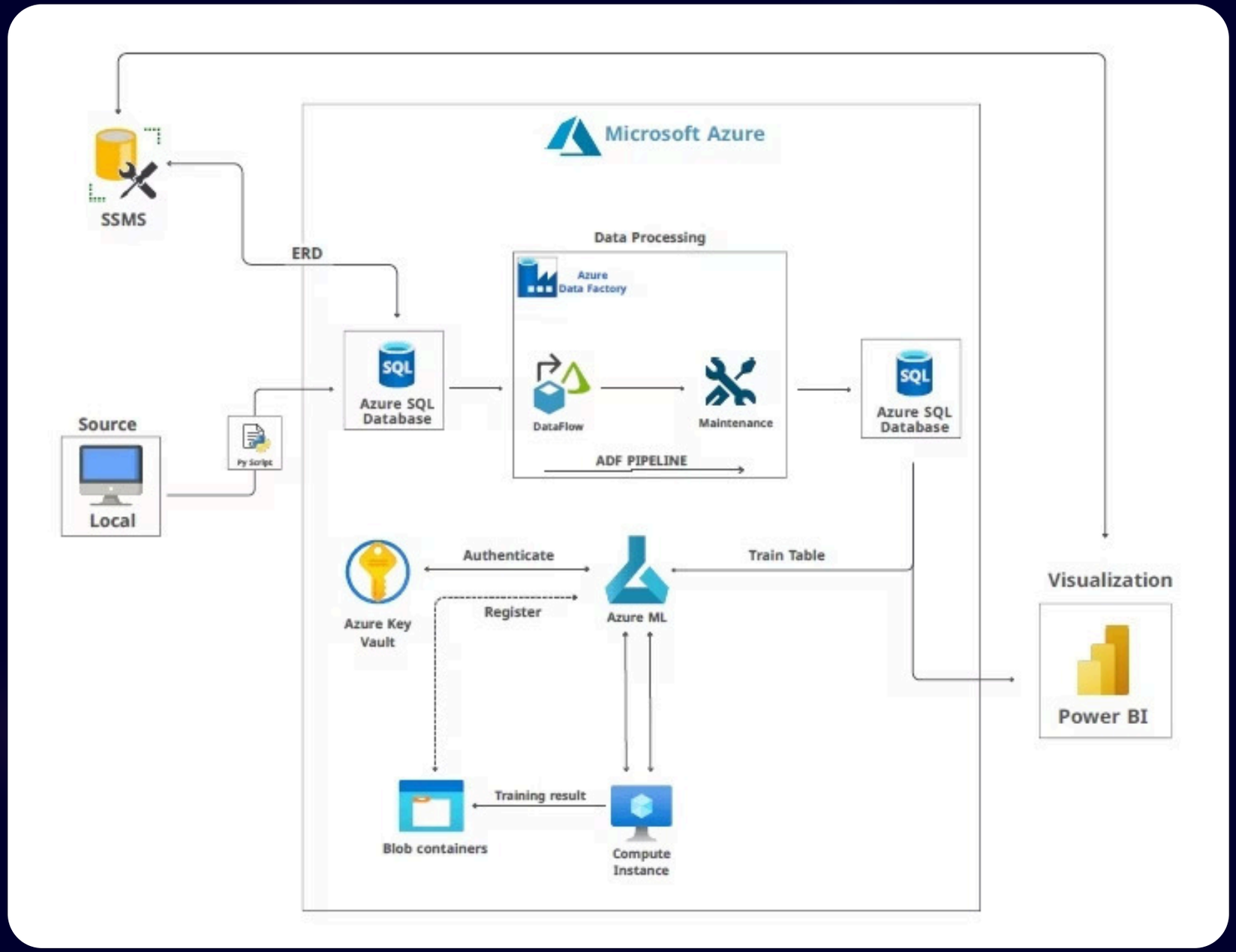
Meat the Team:

- Mai Mamdooh
  - Power BI (Analysis + Dashbord)
  - Presentation
- Naira ELazab
  - Sql Analysis
  - Presentation
- Sayed Elmasry
  - Azure
  - ML
- Verina Fouad
  - SQL
  - Power Bi (dashboard theme)
- Rewan Gamal
  - Power BI (analysis)
- Zaineb Elghoul
  - SQL
- Nermeen Reda

Main Goals:

1. **Identify Top-Performing Properties**
   - Analyze sales and profitability by property type and Location.
2. **Evaluate Agents & Customer Profitability**
3. **Forecast Future Sales**
   - Apply predictive modeling (time series, regression) for demand and revenue trends.
4. **Develop Interactive Dashboard**
   - Create dynamic visualizations for real-time insights and decision-making.

Data WorkFlow:



About the Data Set:

Our dataset consisted of 5 main tables:

- **Sales (Fact Table)**

  Contains sales transactions: SalePrice, Profit, SaleDate

  - Linked to Clients, Agents, Properties

- **Agents** (Dimension Table)

  Info about sales agents: Name, Email, Phone

- **Clients** (Dimension Table)

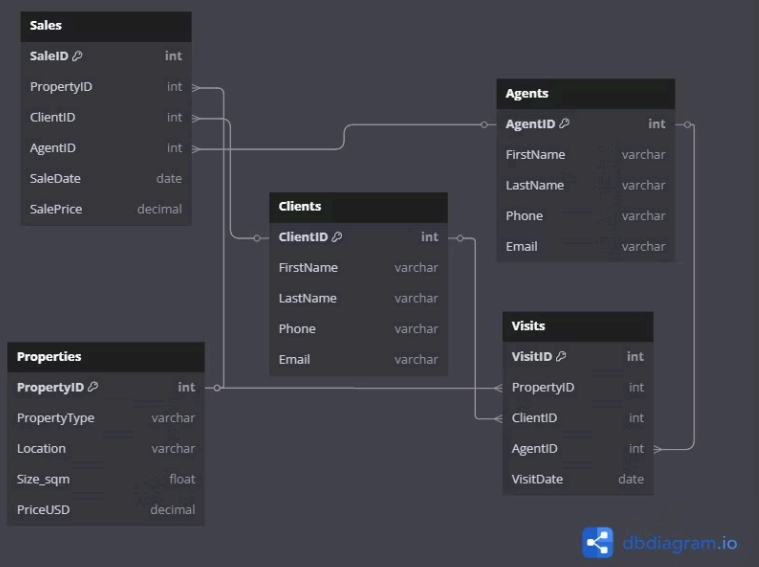
  Info about buyers: Name, Email, Phone

- **Properties** (Dimension Table)

  Details about each property: Location, Type, Price, Size

- **Visits** (Log Table)

  Tracks client visits: Agent, Client, Property, VisitDate



Exploring and Preprocessing

By Azure

Data Factory

Data Flow

In this section, we designed and implemented an integrated data flow using Azure Data Factory (ADF) to prepare the model table that will be used later in the analysis process and build the machine learning model. This flow is a fundamental and vital step in the pre-processing phase, as it collects data from multiple sources, processes it, and transforms it into a unified format suitable for analysis.

- Processing Steps VisitCountAggregate

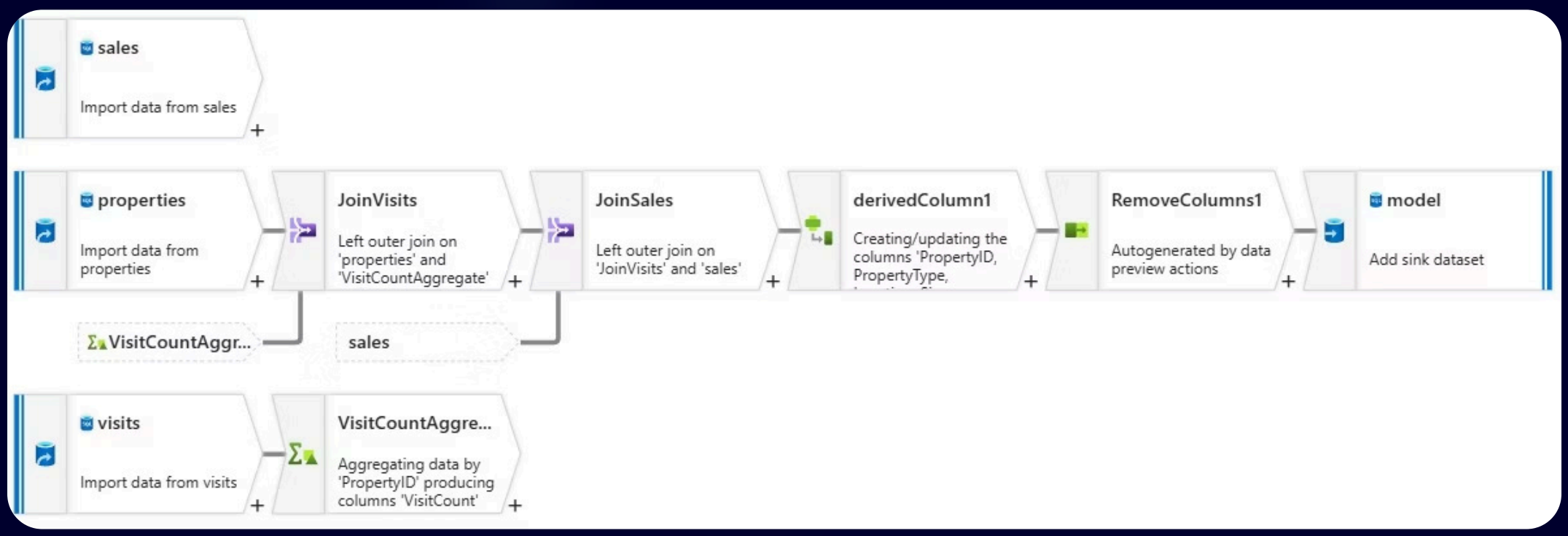
  > Aggregate visit data by PropertyID and generate VisitCount columns JoinVisits: Left outer join between properties data and VisitCountAggregate data JoinSales: Left outer join between the result of the previous join (JoinVisits) and sales data derivedColumn1: Create/update PropertyID and PropertyType columns

- RemoveColumns1

  > Remove unnecessary columns (automatically generated by data preview actions)

- Output model

  > Add the resulting dataset as a data sink to the model



Analysis and Visualization

SQL & Power BI

Sales Performance

Total sales value over time (monthly)

```
SELECT    MONTH(SaleDate) AS MonthNum,
CAST(SUM(SalePrice) AS INT) AS TotalSales
FROM
   Sales
GROUP BY
MONTH(SaleDate)
ORDER BY
    MonthNum DESC;
```

Key Insights

- Strong seasonal pattern with December showing the highest sales (likely year-end closings)
- Mid-year peak in May suggests a second strong selling season
- August shows the lowest performance, indicating a potential summer slowdown

Total sales value over time (quarterly)

```
SELECT    DATEPART(QUARTER, SaleDate) AS QrtNum,
CAST(SUM(SalePrice) AS INT) AS TotalSales
FROM
   Sales
GROUP BY
DATEPART(QUARTER, SaleDate)
ORDER BY
    QrtNum;
```

Key Insights

- Q2 is the strongest quarter for sales, aligning with the monthly data showing May as a peak month
- Q3 shows the lowest performance, confirming the summer slowdown seen in the monthly data

Total sales value over time (yearly)

```
SELECT    YEAR(SaleDate) AS Year,
CAST(SUM(SalePrice) AS INT) AS TotalSales
FROM
   Sales
GROUP BY
YEAR(SaleDate)
ORDER BY
Year;
```

Key Insights

- Significant growth from 2023 to 2024 (67% increase)
- The strong year-over-year growth indicates successful business expansion or market improvement

Average sale value per property type

```
SELECT    p.PropertyType,
CAST(ROUND(AVG(s.SalePrice), 2) AS DECIMAL(18,2)) AS AvgSaleValue
FROM
   Sales s
Inner JOIN
   Properties p
ON s.PropertyID = p.PropertyID
GROUP BY
   p.PropertyType
ORDER BY
   AvgSaleValue DESC;
```

Key Insights

- Warehouses command the highest average sale price, likely reflecting their commercial value and larger size
- Commercial properties (Warehouse, Office, Retail) generally sell for higher prices than residential properties
- The difference between highest and lowest average sale values is only about 6.6%
- Despite Apartments being the most common property type, they have the lowest average sale value

Conversion rate per property

```
WITH PropertyStats AS (    SELECT
       p.PropertyID,
       p.PropertyType,
COUNT(DISTINCT v.VisitID) AS VisitCount,
CASE WHEN
COUNT(DISTINCT s.SaleID) > 0 THEN 1
ELSE 0
END AS WasSold
FROM
       Properties AS p
LEFT JOIN Visits AS v
ON p.PropertyID = v.PropertyID
LEFT JOIN Sales AS s
ON p.PropertyID = s.PropertyID
GROUP BY
       p.PropertyID, p.PropertyType
)
SELECT
   PropertyID,
   PropertyType,
   VisitCount,
   WasSold,
CASE WHEN
       VisitCount = 0 THEN 0
WHEN WasSold = 1 THEN CAST(ROUND(100.0 / VisitCount, 2) AS DECIMAL(5,2))
ELSE 0
END AS ConversionRatePercent
FROM
   PropertyStats
ORDER BY
   ConversionRatePercent DESC;
```

Key Insights

- Properties with 100% conversion rate (sold after first visit) indicate highly desirable assets or effective pricing
- The conversion rate decreases predictably as the number of visits increases

Conversion rate per agent

```
SELECT    a.AgentID,
CONCAT(a.FirstName, ' ', a.LastName) AS AgentName,
COUNT(DISTINCT s.SaleID) AS NumberOfSales,
COUNT(DISTINCT v.VisitID) AS NumberOfVisits,
CAST(ROUND((COUNT(DISTINCT s.SaleID) * 100.0) / NULLIF(COUNT(DISTINCT v.VisitID), 0),2) AS DECIMAL(5,2)) AS
ConversionRatePercentage
FROM
   Agents AS a
LEFT JOIN
   Sales AS s
ON a.AgentID = s.AgentID
LEFT JOIN
   Visits AS v
ON a.AgentID = v.AgentID
GROUP BY
   a.AgentID, a.FirstName, a.LastName
ORDER BY
   ConversionRatePercentage DESC;
```

Key Insights

- Top-performing agents achieve significantly higher conversion rates than others
- Conversion rate is a better performance indicator than raw sales numbers as it measures efficiency

Time on market before sale

```
SELECT
   s.PropertyID,
   p.PropertyType,
MIN(v.VisitDate) AS FirstVisitDate,
   s.SaleDate,
DATEDIFF(DAY, MIN(v.VisitDate), s.SaleDate) AS DaysOnMarket
FROM
   Sales AS s
Inner JOIN
   Properties AS p
ON s.PropertyID = p.PropertyID
Inner JOIN
   Visits AS v
ON s.PropertyID = v.PropertyID
GROUP BY
   s.PropertyID, p.PropertyType, s.SaleDate
ORDER BY
   DaysOnMarket DESC;
```

Key Insights

- Properties show varying time periods between first visit and sale
- Some properties sell very quickly after the first visit

Key Sales Metrics

Total Clients

1500

Total Properties

1000

Total Visits

5000

Total Sales

$1.54B

Avg Property Price

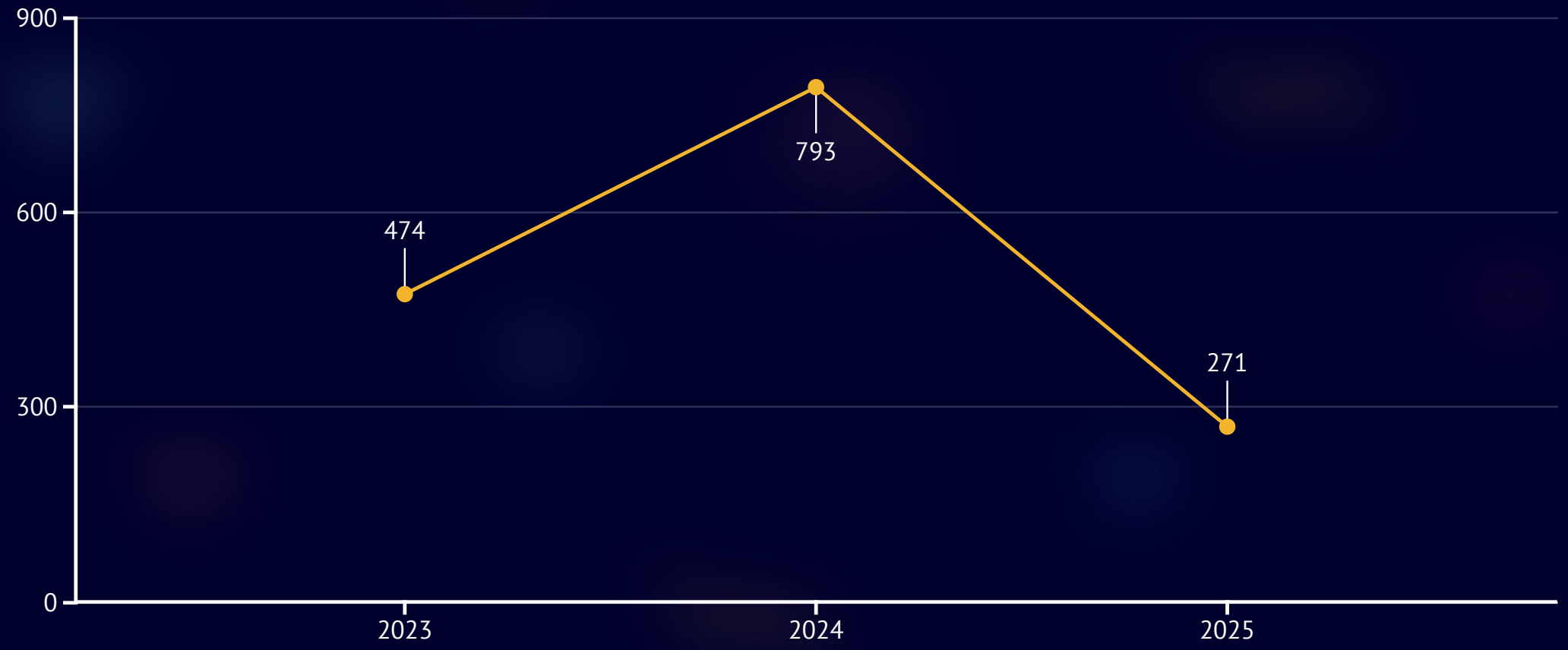
$521K

Conversion Rate

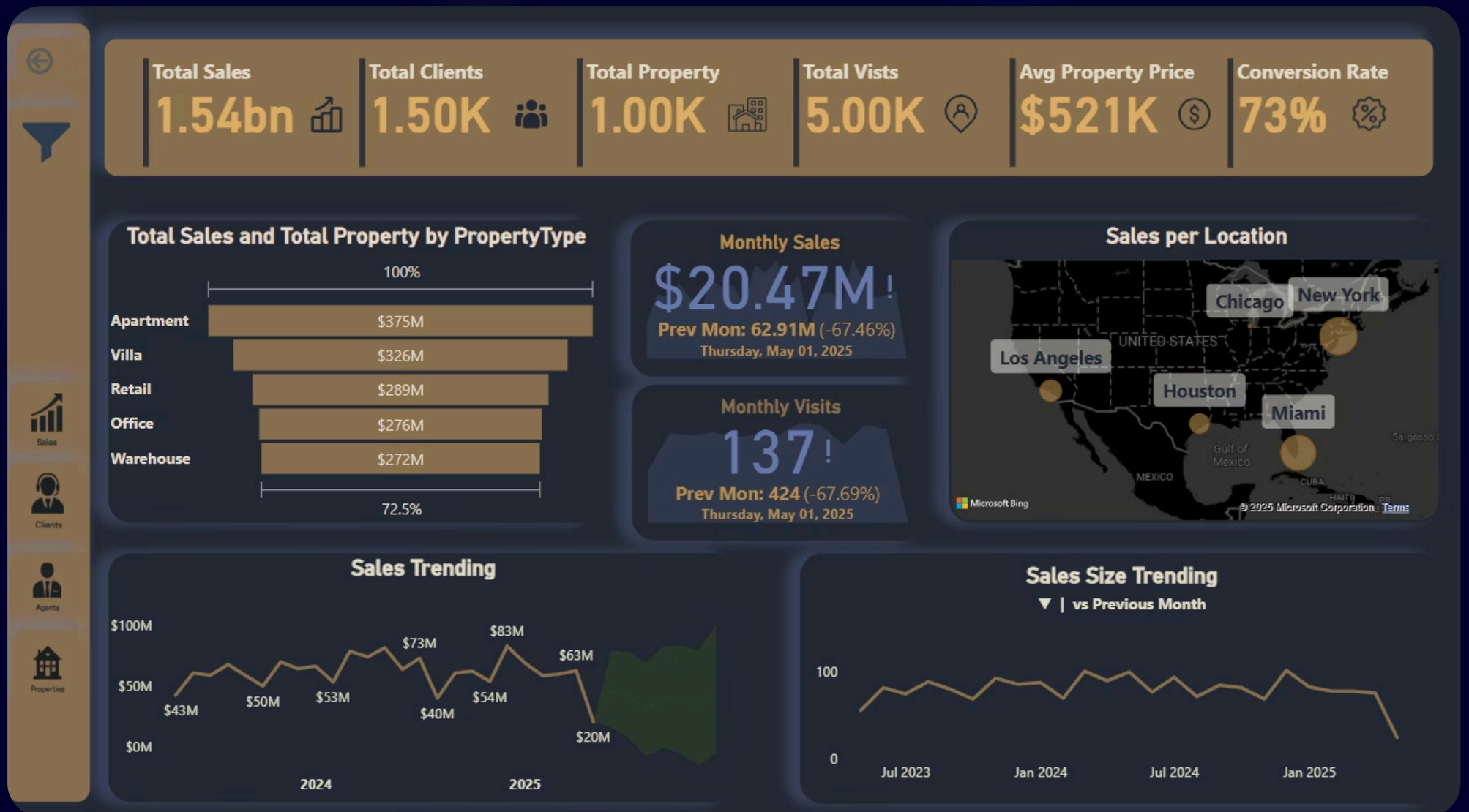
73%

Sales Trending

✓ Past sales are stable with slight fluctuations, but there's a dip in early 2025.



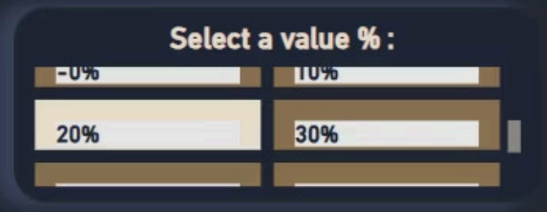
Sales Dashbord

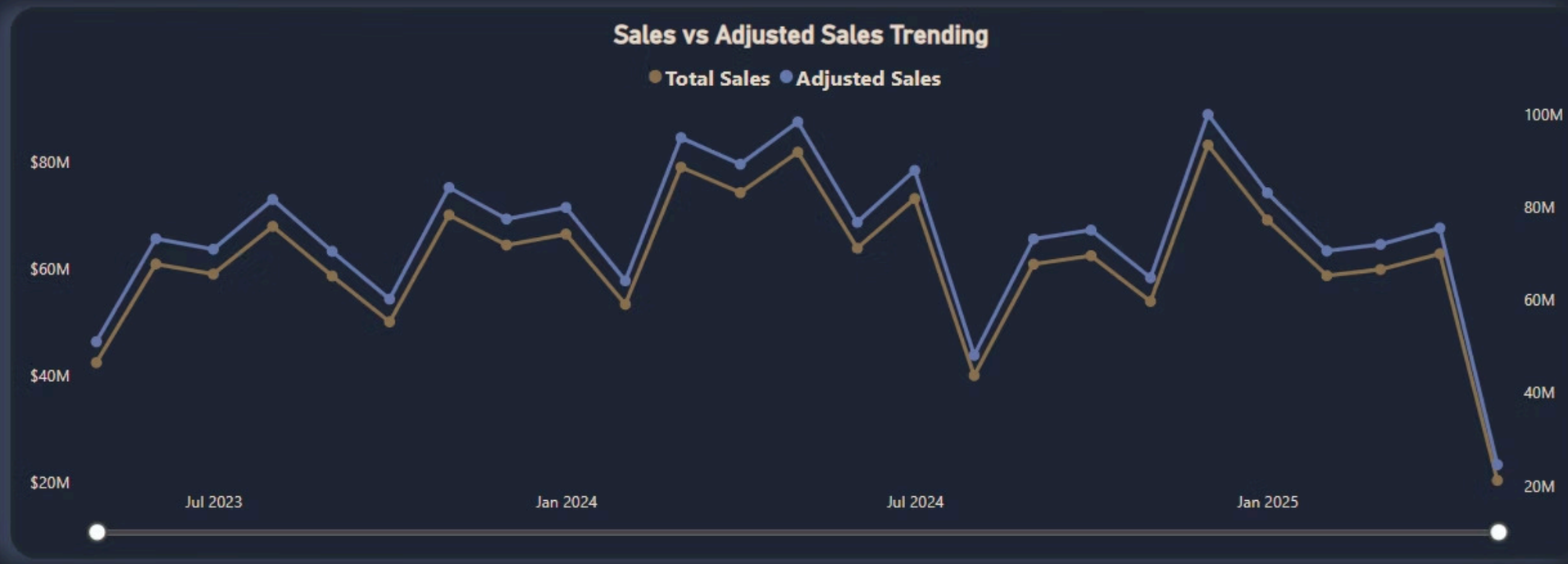


Sales Adjustment

Useful for forecasting potential impacts of various discount strategies on total Sales.

*A **current sales performance** and offers a **'what if' simulation** of how price adjustments could impact our future sales:*





Agent Performance

Number of sales per agent

```
SELECT
    a.AgentID,
    CONCAT(a.FirstName, ' ', a.LastName) AS AgentName,
    COUNT(s.SaleID) AS NumberOfSales
FROM
    Agents AS a
LEFT JOIN
    Sales AS s
    ON a.AgentID = s.AgentID
GROUP BY
    a.AgentID, a.FirstName, a.LastName
ORDER BY
    NumberOfSales DESC;
```

Key Insights

- Clear performance tiers emerge among agents
- Sales counts vary significantly across agents
- Top-performing agents handle substantially more sales than others

Number of client visits per agent

```
SELECT
    a.AgentID,
    CONCAT(a.FirstName, ' ', a.LastName) AS AgentName,
    COUNT(v.VisitID) AS NumberOfClientVisits
FROM
    Agents AS a
LEFT JOIN
    Visits AS v
    ON a.AgentID = v.AgentID
GROUP BY
    a.AgentID, a.FirstName, a.LastName
ORDER BY
    NumberOfClientVisits DESC;
```

Key Insights

- High visit counts indicate active prospecting and client engagement
- Some agents may have high visit counts but lower sales, suggesting conversion challenges, only Samantha Vargas was of the top Sales
- Activity levels don't always correlate directly with sales performance

Conversion rate per agent (percentage)

```
SELECT
    a.AgentID,
    CONCAT(a.FirstName, ' ', a.LastName) AS AgentName,
    COUNT(DISTINCT s.SaleID) AS NumberOfSales,
    COUNT(DISTINCT v.VisitID) AS NumberOfVisits,
    CAST(ROUND((COUNT(DISTINCT s.SaleID) * 100.0) / NULLIF(COUNT(DISTINCT v.VisitID), 0),2) AS DECIMAL(5,2))
AS ConversionRatePercentage
FROM
    Agents AS a
LEFT JOIN
    Sales AS s
        ON a.AgentID = s.AgentID
LEFT JOIN
    Visits AS v
        ON a.AgentID = v.AgentID
GROUP BY
    a.AgentID, a.FirstName, a.LastName
ORDER BY
    ConversionRatePercentage DESC;
```

Key Insights

- Conversion efficiency varies significantly across the team
- Some agents may have fewer total sales but higher conversion efficiency
- This identifies different sales approaches and effectiveness

Average sale value handled by each agent

```
SELECT
    a.AgentID,
    CONCAT(a.FirstName, ' ', a.LastName) AS AgentName,
    CAST(ROUND(AVG(s.SalePrice), 2) AS DECIMAL(15,2)) AS AverageSaleValue
FROM
    Agents AS a
LEFT JOIN
    Sales AS s
        ON a.AgentID = s.AgentID
GROUP BY
    a.AgentID, a.FirstName, a.LastName
ORDER BY
    AverageSaleValue DESC;
```

Key Insights

- Highest AverageSaleValue Agent is Carolyn Terry
- Average transaction values vary by agent
- This helps in strategic assignment of listings to appropriate agents

Agent Performance

Total Agents

100

Avg Agent Sales

$15. 39M

Avg Days to Sell

72 Day

Top Agent

Samantha Vargas

$24.22M in sales and 27 sold properties.

Top 10 Agent by Sales

Price Status

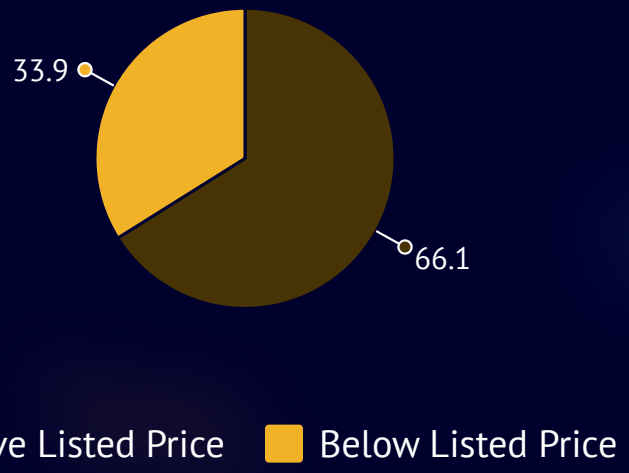
Illustrates the price difference between the listed price and the actual selling price of properties — highlighting the agent's ability to negotiate a higher deal. It shows the profit margin achieved by agents

Price Negotiation

66.1% of properties sold above listed price.

Price Differences

Significant price gaps highlight negotiation skills and risks.



Agent Dashbord

Top 10 Agents

Agent_Name	Sales Price	Actual Price	Price diff	#Clients	#Sold Property
Cody Ramsey	$20.50M	14.32M	$6.17M	25	26
Francisco Williams	$24.18M	13.48M	$10.69M	29	29
Gordon Wilson	$23.28M	12.98M	$10.31M	28	28
Laura Martin	$20.85M	12.21M	$8.64M	25	24
Mark Lopez	$20.99M	9.25M	$11.73M	24	24
Michelle Davis	$20.64M	14.39M	$6.25M	28	28
Robert Miller	$21.44M	7.81M	$13.63M	23	23
Samantha Vargas	$24.22M	16.65M	$7.57M	28	27
Tammy Cohen	$20.40M	13.37M	$7.03M	22	22
Tammy Walker	$20.45M	11.63M	$8.81M	23	23

Client Engagement

Number of properties visited per client

```
SELECT
   c.ClientID,
   CONCAT(c.FirstName, ' ', c.LastName) AS ClientName,
   COUNT(DISTINCT v.PropertyID) AS NumberOfPropertiesVisited
FROM
   Clients AS c
LEFT JOIN
   Visits AS v
   ON c.ClientID = v.ClientID
GROUP BY
   c.ClientID, c.FirstName, c.LastName
ORDER BY
   NumberOfPropertiesVisited DESC;
```

Key Insights

- Clients with many visits without purchases may need more targeted recommendations
- Clients who purchase after few visits represent efficient transactions
- This metric helps prioritize follow-up with highly engaged prospects

Top clients by sale value

```
SELECT TOP 10
   c.ClientID,
   CONCAT(c.FirstName, ' ', c.LastName) AS ClientName,
   CAST(SUM(s.SalePrice) AS FLOAT) AS SaleValue
FROM
   Clients c
Inner JOIN
   Sales s ON c.ClientID = s.ClientID
GROUP BY
   c.ClientID, c.FirstName, c.LastName
ORDER BY
   SaleValue DESC;
```

Key Insights

- Significant variation in spending among top clients
- Understanding characteristics of top clients can inform acquisition strategies
- Potential for developing targeted loyalty programs for high-value clients

First-time vs repeat buyers

```
WITH ClientSaleCounts AS (
   SELECT
      ClientID,
      COUNT(SaleID) AS NumberOfSales
   FROM
      Sales
   GROUP BY
      ClientID
)
SELECT
   c.ClientID,
   CONCAT(c.FirstName, ' ', c.LastName) AS ClientName,
      ISNULL(csc.NumberOfSales, 0) AS TotalSalesMade,
   CASE WHEN
      csc.NumberOfSales = 1 THEN 'First-Time Buyer'
      WHEN csc.NumberOfSales > 1 THEN 'Repeat Buyer'
      ELSE 'Not a Buyer'
   END AS BuyerType
FROM
   Clients AS c
LEFT JOIN
   ClientSaleCounts AS csc
      ON c.ClientID = csc.ClientID
ORDER BY
   c.LastName, c.FirstName;
```

Key Insights

- Clients categorized as First-Time Buyers, Repeat Buyers, or Not a Buyer
- Distribution shows the mix of new vs. returning customers

Region-based client interest

```
SELECT
   c.ClientID,
   CONCAT(c.FirstName, ' ', c.LastName) AS ClientName,
   p.Location AS City,
   COUNT(v.VisitID) AS NumberOfVisitsInCity
FROM
   Clients AS c
Inner JOIN
   Visits AS v ON
   c.ClientID = v.ClientID
Inner JOIN
   Properties AS p
   ON v.PropertyID = p.PropertyID
GROUP BY
   c.ClientID, c.FirstName, c.LastName, p.Location
ORDER BY
   NumberOfVisitsInCity DESC;
```

Key Insights

- This data enables personalized property recommendations by location
- Marketing can be targeted to promote properties in preferred locations
- Many clients show strong preferences for specific locations(Chicago,New York)

Client Analysis

Total Clients

1500

Active Clients

1096

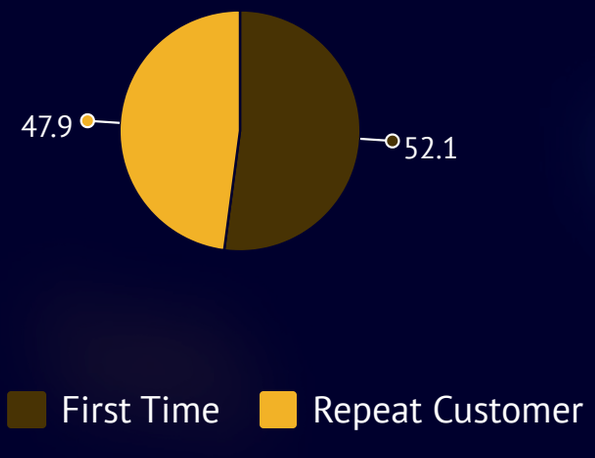
Total Visits

5000

4 visits per client.

Type

More than half of the clients are returning, which represent strong customer satisfaction and long-term relationships.



Top Clients

- **Laura Allen** generated the highest sales ($7.03M) and highest profit ($3.52M) with 13 visits.
- **Kimberly Bowen** is one of the top 10 although it has negative profit (-$0.22M), suggesting either a miss pricing or a strategic discount.

Top 10 Client

Client Name	Sales	Profit	#Property	#Vists
Laura Allen	$7.03M	$3.52M	6	13
William Brown	$6.08M	$2.75M	6	14
Miranda Gomez	$5.31M	$2.26M	7	5
Deborah Gibson	$5.06M	$2.89M	5	2
Brenda Cobb	$4.98M	$1.86M	5	5
Michael Martinez	$4.91M	$2.37M	5	6
Kenneth Smith	$4.87M	$2.34M	5	8
James Andrade	$4.73M	$3.71M	4	4
Michelle Martin	$4.67M	$1.02M	5	5
Kimberly Bowen	$4.50M	($0.22M)	6	4

Client Dashbord

Property Analytics

Number of listed properties by type and location

```
SELECT
   PropertyType,
Location,
COUNT(*) AS NumberOfProperties
From
   Properties
Group By
   PropertyType, Location
ORDER BY
   PropertyType, Location;
```

Key Insights

- Miami leads in both Apartments (55) and Villas (52), suggesting a strong residential market
- New York has the highest concentration of Office properties (52)
- Chicago has the strongest retail property presence (46)
- Los Angeles leads in warehouse properties (46)

Average price per square meter per city

```
SELECT
   Location AS City,
Round(AVG(PriceUSD / NULLIF(Size_sqm, 0)),2) AS AvgPricePerSqm
FROM
   Properties
GROUP BY
Location
ORDER BY
   AvgPricePerSqm DESC;
```

Key Insights

- Chicago has the highest property value per square meter, despite not having the highest total property count
- Surprisingly, New York ranks only 4th in price per square meter despite its reputation for expensive real estate

Distribution of property types

```
SELECT
   PropertyType,
COUNT(*) AS NumberOfProperties
FROM
   Properties
GROUP BY
   PropertyType
ORDER BY
   NumberOfProperties DESC;
```

Key Insights

- Residential properties (Apartments and Villas combined) make up the largest segment of the portfolio (423 properties, ~42%)
- Commercial properties (Retail, Warehouse, Office) collectively represent the majority of the portfolio (577 properties, ~58%)
- Apartments are the most common property type, suggesting a focus on residential rental or multi-family investments

Top 10 most expensive properties

```
SELECT    TOP 10
   PropertyID,
   PropertyType,
   PriceUSD
FROM
   Properties
ORDER BY
   PriceUSD DESC;
```

Key Insights

- The most expensive property is an Apartment (ID: 297) valued at $998,279

Top 10 most visited properties

```
SELECT
   TOP 10
   v.PropertyID,
   p.PropertyType,
COUNT(*) AS NumberOfVisits
FROM
   Visits AS v
Inner JOIN
   Properties AS p
ON v.PropertyID = p.PropertyID
GROUP BY
   v.PropertyID, p.PropertyType, p.Location
ORDER BY
   NumberOfVisits DESC;
```

Key Insights

- Property ID 353 (Apartment) is the most visited with 13 visits
- Apartments and Warehouses generate the most client interest

Properties Overview

Total Properties

1000

Sold Properties

863 (86.3% sold rate)

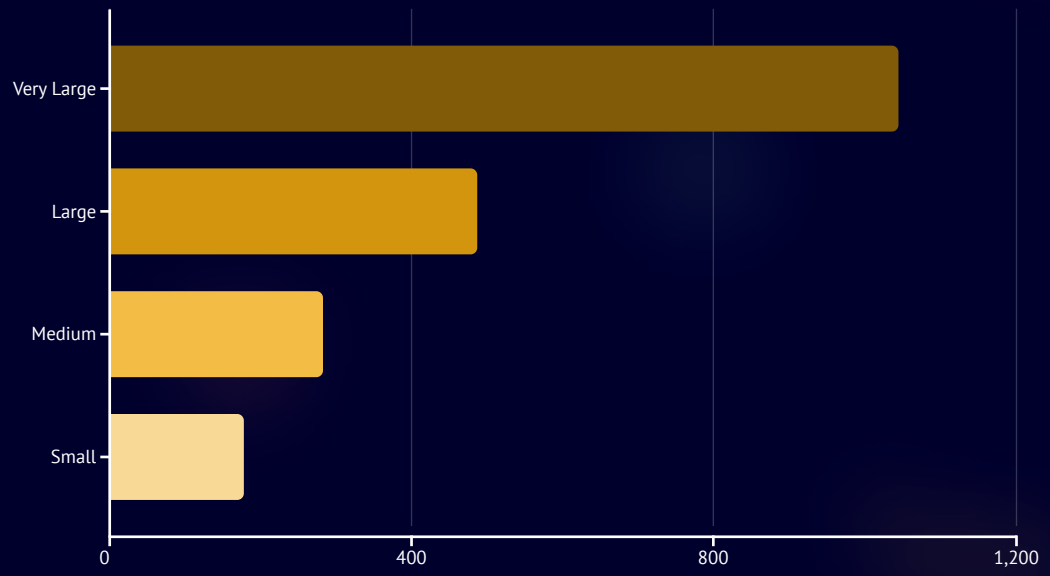
Avg Property Size

266.8 sqm

Average Price per sqm

$1950

Property Size



- Small: Properties up to 90 sqm.
- Medium: Properties from 91 sqm to 150 sqm.
- Large: Properties from 151 sqm to 250 sqm.
- Very Large: Properties over 250 sqm.

Top 10 Most Visited Properties

- Locations: Mostly in Chicago and Miami
- Top Types: Apartments, Retail, Offices, and Warehouses
- Observation: Properties priced between $180K–$700K get the highest attention, especially medium to very large sizes.

Top 10 Properties by visits

PropertyID	Location	Type	Price	#Vists	Size
353	Chicago	Apartment	188K	13	Medium
380	Chicago	Retail	707K	12	Very Large
398	Miami	Apartment	189K	12	Large
652	Miami	Office	839K	12	Very Large
821	Miami	Apartment	692K	12	Large
150	New York	Villa	550K	11	Very Large
194	Houston	Warehouse	89K	11	Very Large
392	Chicago	Office	767K	11	Medium
447	Miami	Warehouse	260K	11	Very Large
517	Chicago	Warehouse	329K	11	Very Large
607	Chicago	Office	130K	11	Medium
714	Los Angeles	Retail	607K	11	Very Large
853	Houston	Warehouse	227K	11	Very Large
876	Los Angeles	Villa	751K	11	Very Large
924	Miami	Retail	613K	11	Very Large
954	Chicago	Apartment	246K	11	Large

Properties Dashboard

Location-Based Insights

Sales heatmap by city (volume)

```
SELECT
   p.Location,
   COUNT(s.SaleID) AS SalesNum
FROM
   Properties AS p
INNER JOIN
   Sales AS s
   ON p.PropertyID = s.PropertyID
GROUP BY
   p.Location
ORDER BY
   SalesNum DESC;
```

Key Insights

- Sales volume varies by location
- High-volume areas may indicate strong market demand or effective local marketing (New York : 431 SalesNum / Miami: 423 SalesNum)

Sales heatmap by city (value)

```
SELECT
   p.Location,
      CAST(SUM(s.SalePrice) AS FLOAT) AS SaleValue
FROM
   Properties p
Inner JOIN
   Sales AS s
      ON p.PropertyID = s.PropertyID
GROUP BY
   p.Location
ORDER BY
   SaleValue DESC;
```

Key Insights

- Total sales value by location Top (New York , Miami)
- Financial significance of different markets

High-performing areas (most sold)

```
SELECT
   p.Location,
   COUNT(s.SaleID) AS SalesNum
FROM
   Properties AS p
INNER JOIN
   Sales AS s
   ON p.PropertyID = s.PropertyID
GROUP BY
   p.Location
ORDER BY
   SalesNum DESC;
```

Key Insights

- High-performing areas represent locations with strong demand (New York : 431 SalesNum / Miami: 423 SalesNum)
- These areas represent locations where properties move quickly
- Understanding high-performing areas helps focus marketing and acquisition

Average visit-to-sale ratio per location

```
WITH SoldProperties AS (
   SELECT
      p.PropertyID,
      p.Location
   FROM
      Properties p
      INNER JOIN Sales AS s
                  ON p.PropertyID = s.PropertyID
),
VisitsPerSoldProperty AS (
   SELECT
      sp.PropertyID,
      sp.Location,
      COUNT(v.VisitID) AS VisitsNum
   FROM
      SoldProperties AS sp
      LEFT JOIN Visits AS v
            ON sp.PropertyID = v.PropertyID
   GROUP BY
      sp.PropertyID, sp.Location
)
SELECT
   Location,
   AVG(VisitsNum * 1.0) AS AvgVisitToSaleRatio
FROM
   VisitsPerSoldProperty
GROUP BY
   Location
ORDER BY
   AvgVisitToSaleRatio DESC;
```

Key Insights

- Some areas may have higher visit-to-sale ratios, indicating more efficient markets
- Areas with low ratios may need pricing adjustments or improved marketing
- Chicago : 12.391566 AvgVisitToSaleRatio
- New York : 12.011173 AvgVisitToSaleRatio
- Miami: 11.966480 AvgVisitToSaleRatio
- Los Angeles : 11.339181 AvgVisitToSaleRatio
- Houston: 10.898809 AvgVisitToSaleRatio

Location Analysis

New York

☒ Price per sqm: $1,930

☒ Most Sold: Apartment (23.24%)

Why Apartments?

• High population density and limited land space make vertical housing (apartments) the most practical.

• Many professionals and international buyers prefer smaller, investment-friendly units.

Los Angeles

☒ Balanced average sqm price: $1,960

☒ Most Sold: Apartment (25.92%)

Why Apartments?

• LA has soaring real estate prices, making apartments more affordable for both buyers and renters.

• Apartments cater to a wide demographic — students, young professionals, and newcomers.

Houston

☒ Price per sqm: $1,948

☒ Most Sold: Warehouse (22.83%)

Why Warehouses?

• Houston is a logistics and industrial powerhouse, home to one of the largest ports in the U.S.

• Strong presence of oil & gas, manufacturing, and shipping industries increases warehouse demand.

• The rise of e-commerce fuels high demand for storage and fulfillment centers

Chicago

☒ Highest price per sqm: $2,040

☒ Most Sold: Villa (22.46%)

Why Villas?

- Chicago offers more land and lower property prices compared to coastal cities, enabling larger homes.
- Many middle-to-upper class residents prefer spacious villas in suburbs like Naperville or Evanston.
- Strong demand from families looking for multi-bedroom, detached homes with outdoor space.
- Low-rise housing is common, making villas more accessible.

Chicago is most visited and top-selling location.

Miami

☒ Lowest price per sqm: $1,900

☒ Most Sold: Retail (28.85%)

Why Retail?

• Miami is a tourism and shopping hub with massive foot traffic from locals and tourists.

• Heavy commercial zones, especially in places like Brickell, Wynwood, and South Beach, drive demand for retail spaces.

• The Latin American market connection makes Miami a hotspot for international retail brands.

Milestone Queries

Potential Customers for Email Campaigns

```
WITH ClientSaleCounts AS (
    SELECT
        ClientID,
        COUNT(SaleID) AS NumberOfSales
    FROM
        Sales
    GROUP BY
        ClientID
)
SELECT
    c.ClientID,
    CONCAT(c.FirstName, ' ', c.LastName) AS ClientName,
    c.Email,
    ISNULL(csc.NumberOfSales, 0) AS TotalSalesMade
FROM
    Clients AS c
LEFT JOIN
        ClientSaleCounts AS csc
        ON c.ClientID = csc.ClientID
WHERE
    csc.NumberOfSales = 1
        OR
        csc.NumberOfSales IS NULL
ORDER BY
    c.LastName, c.FirstName;
```

Key Insights

- First-time buyers (TotalSalesMade = 1): These clients have already completed one transaction and represent opportunities for repeat business
- Prospects (TotalSalesMade = 0): These are registered clients who haven't yet made a purchase
- Email campaigns can be customized differently for each segment:
  - First-time buyers: Retention-focused messaging, loyalty programs, complementary properties
  - Prospects: Conversion-focused messaging, introductory offers, high-visibility properties

Not Visited Properties to Focus On for Campaigns and Offers

```
SELECT
    p.PropertyID,
        p.PropertyType,
        p.Location
FROM
    Properties AS p
LEFT JOIN
    Visits AS v
    ON p.PropertyID= v.PropertyID
WHERE
    v.VisitID IS NULL
```

Key Insights

- Property types include: Warehouse (2), Retail (4), Office (2), Apartment (1)
- Locations include: New York, Chicago, Houston, Miami
- These 10 properties represent inventory that is generating zero interest, indicating potential issues with:
  - Inadequate marketing visibility
  - Pricing misalignment with market expectations
  - Property features that don't match current demand
  - Possible location or accessibility challenges
- The distribution across property types suggests the issue isn't limited to a specific category
- These properties should be prioritized for:
  - Special promotional offers or pricing adjustments
  - Enhanced visual content (professional photography, virtual tours)
  - Agent incentives for showing these properties
- The relatively small number (10 out of 1000 properties) indicates generally good inventory exposure

Month-over-Month (MoM) Percentage Change in Sales

```
SELECT
    YEAR(SaleDate) As Year,
    MONTH(SaleDate) AS Month,
    SUM(SalePrice) AS Total_Sales,
    ROUND(
            (SUM(SalePrice) -
                        LAG(SUM(SalePrice)) OVER (ORDER BY YEAR(SaleDate), MONTH(SaleDate)) ) * 100.0 /
                        NULLIF(LAG(SUM(SalePrice)) OVER (ORDER BY YEAR(SaleDate), MONTH(SaleDate)), 0)
        , 2) AS MOM_per
FROM
    Sales
GROUP BY
    YEAR(SaleDate),
    MONTH(SaleDate)
ORDER BY
    YEAR(SaleDate),
    MONTH(SaleDate);
```

Key Insights

- Consistent end-of-year strength (December shows strong performance)
- Mid-year fluctuations with both strong gains and significant drops
- August 2024 showed a severe drop (-45.24%), followed by strong recovery in September (+51.98%)
- May 2025 shows an alarming decline (-67.46%), requiring immediate investigation
- Seasonal factors strongly influence the real estate market

Year-over-Year (YoY) Percentage Change in Sales

```
SELECT
    YEAR(SaleDate) AS Year,
    SUM(SalePrice) AS Total_Sales,
    ROUND(
        (SUM(SalePrice) -
                    LAG(SUM(SalePrice)) OVER (ORDER BY YEAR(SaleDate))) * 100.0 /
                    NULLIF(LAG(SUM(SalePrice)) OVER (ORDER BY YEAR(SaleDate)), 0)
        , 2) AS YOY_per
FROM
    Sales
GROUP BY
    YEAR(SaleDate)
ORDER BY
    YEAR(SaleDate)
```

Key Insights

- 2023: $474,210,304 (baseline year)
- 2024: $793,376,420 (67.30% increase from 2023)
- 2025: $271,362,351 (-65.80% decrease from 2024, partial year)

Forcasting & Predictive Model

By Azure

Machine Learning Using Azure ML Studio

Setting up the work environment

Create a Compute Instance to run the test and training tasks.

Data Preparation (Preprocessing Pipeline)

Encoding: Converting text values to numbers using appropriate encoding techniques.

Scaling: Standardizing numeric values to improve model performance.

SMOTE: Addressing data imbalances by oversampling minorities.

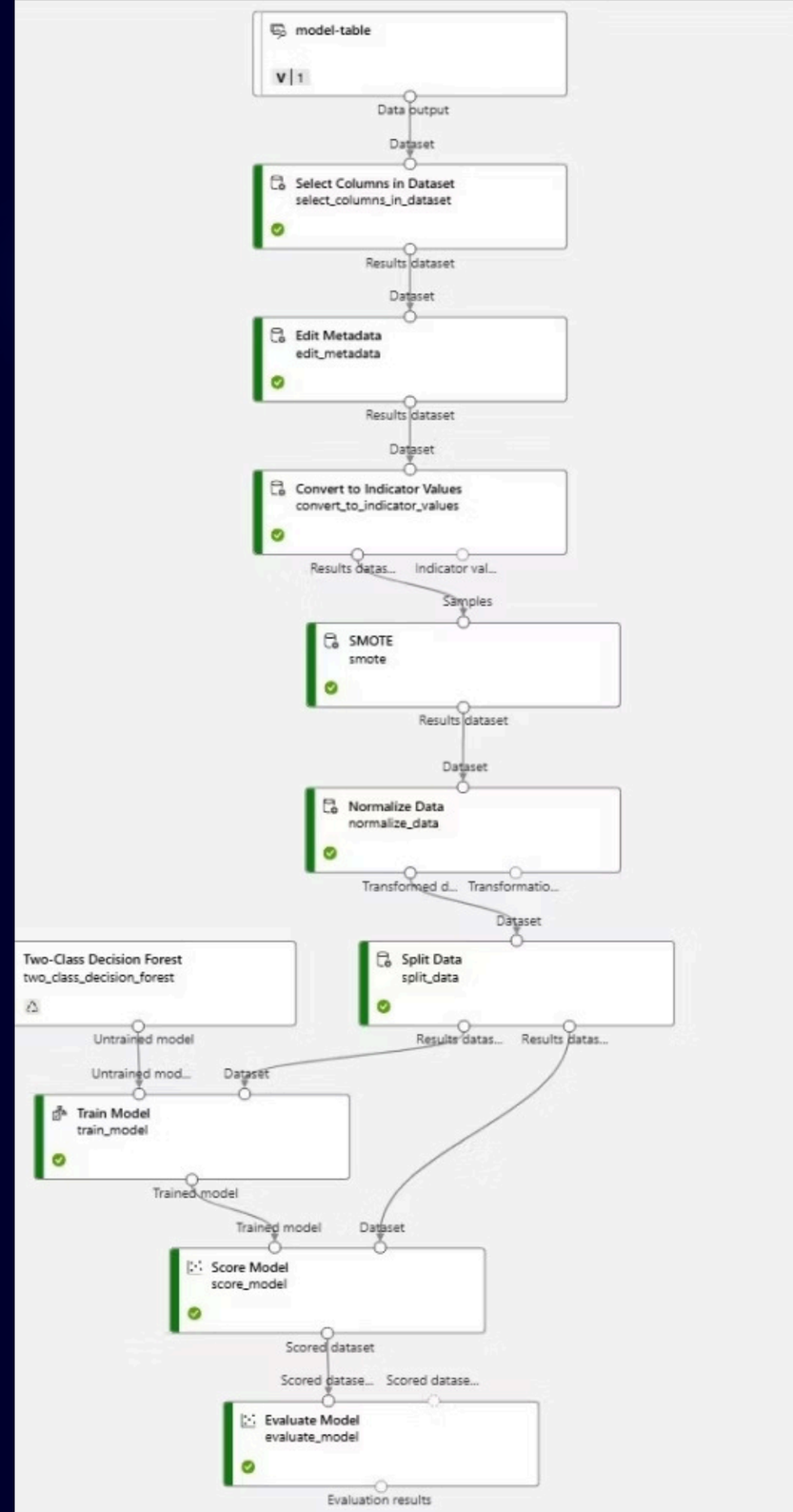
Model Training

Using the Two Class Decision Forest algorithm.

Segmenting the data into training and test data.

Model Evaluation

Calculating key metrics: Accuracy, Precision, Recall, and F1-score.



Model Evaluation Results

The model's performance was evaluated using several criteria, and the results are as follows:

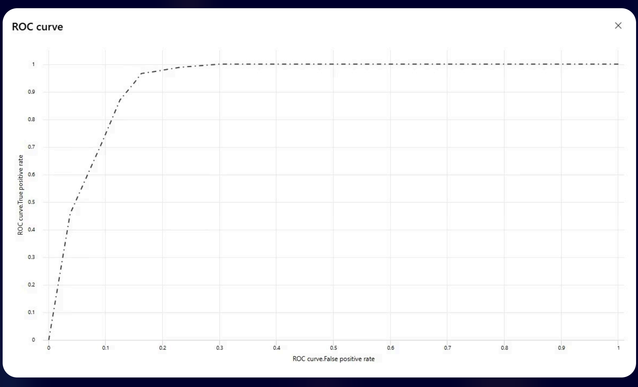
- Key Weighted Metrics
  - Precision: 0.9050445 (90.5%)
  - F1 Score: 0.9144385 (91.4%)
  - Recall: 0.9661017 (96.6%)
  - Area Under the ROC (AUC): 0.9363701 (93.6%)
  - Precision: 0.8680203 (86.8%)

Accuracy	F1 Score	Recall
0.9050445	0.9144385	0.9661017

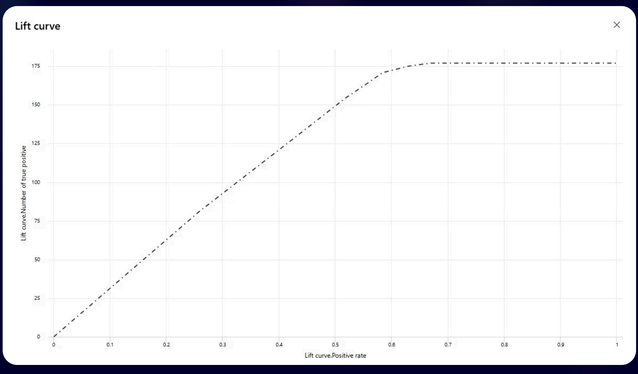
AUC	Precision
0.9363701	0.8680203

- Evaluation Curves

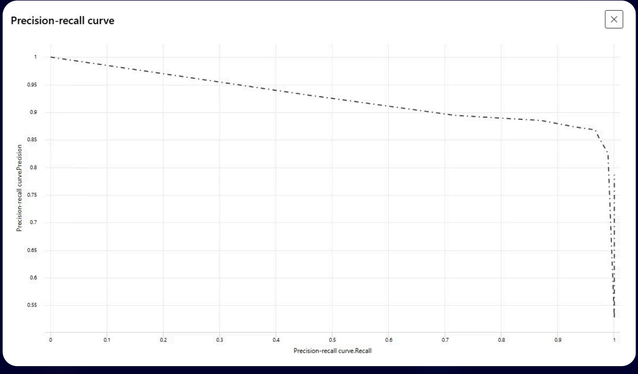
1- ROC curve: Shows excellent model performance with a rapid increase in the growth rate of real drivers versus the growth rate of false drivers, indicating high discrimination ability. For the model.



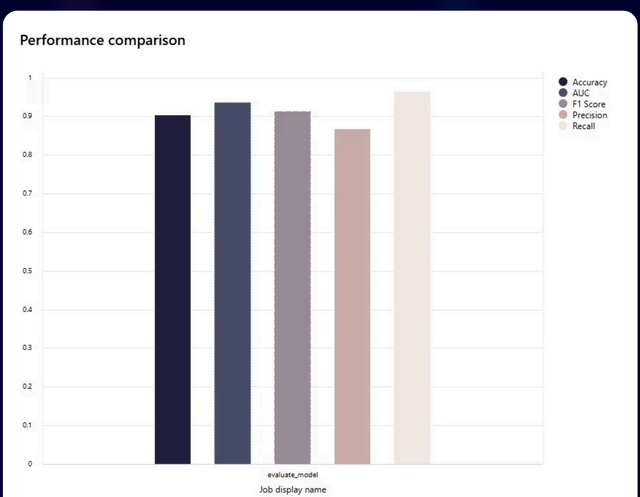
2- Lift curve: Shows that the model performs well compared to the random prediction, with the lift value reaching approximately 175 at 60%.



3- Precision-Recall curve: The kitchen shows a good balance between precision and recall, maintaining high precision values even as recall increases.



- Performance comparison
  - The highest value achieved was for recall at 96.6%, indicating the model's ability to identify positive cases.
  - The lowest value was for precision at 86.8%, indicating that there are some false positives in the model's predictions.
  - The F1 project (91.4%) indicates that the battery is good between precision and recall.



Overall, the model demonstrates excellent performance on the binary classification task, with high cross-class training capability, and the kitchen performs well across various performance benchmarks.

Conclusion & Recommendation

- Market & Product Focus
  - Urban Apartments: Prioritize high-demand cities (NY, LA, Chicago, Miami) for apartments/villas.
  - Large Properties: Expand investment in top-selling categories (e.g., luxury villas, mega-apartments).
  - Warehouses: Leverage Houston's logistics boom; target commercial clients needing storage.
  - Low-Performing Segments: Re-evaluate offices/warehouses—revamp offerings or phase them out.

- Pricing & Profitability
  - Price Discipline:

    - Enforce minimum sale prices (e.g., 90% of market value) to prevent underpricing.

    - Investigate agents with large price gaps (e.g., Kimberly Bowen) to avoid loss-making deals.

  - Negative Profit Cases: Audit transactions with significant deviations to identify root causes.

- Client Management

  • Repeat Clients (52% Segment):

  -Launch loyalty programs (tiered rewards, VIP services, referral incentives).

  -Offer non-price perks (free staging, premium listings) to retain price-sensitive clients.

  • High-Value Clients: Assign top agents to nurture relationships and upsell.

- Agent Performance

  Top Performers (e.g., Samantha Vargas, Mark Lopez):

  - Reward & Replicate: Study their techniques; expand their roles or mentor others.
  - Align with Expertise: Assign agents to cities/property types they excel in (e.g., warehouse specialists to Houston).

  Mid-Tier Agents:

  - Provide targeted training (negotiation, client insights) shadowing top agents.
  - Implement performance benchmarks tied to incentives.

- Operational Efficiency

  Declining Sales/Visits:

  - Investigate causes (market shifts? seasonality? operational gaps?).
  - Boost marketing (geo-targeted campaigns) and client follow-ups.

  Forecasting: Adjust budgets/projections if low sales persist.

- Marketing & Outreach

  Segmented Campaigns:

  - Apartments in NY/LA; retail/warehouses in Miami/Houston.
  - Highlight unique selling points (e.g., "Maximize value, not just speed").

  Transparency: Publish quarterly market reports to justify pricing.

Next Steps

1

Seasonal Campaign

July-August + benefits with local service providers.

2

For properties with extended market times (over 180 days)

systematic review process.

3

Loyalty Program:

Bronze / Silver / Gold.

☆

Value-Added Service Bundling:

complimentary moving services, home warranty packages

Thank You

We appreciate your time and attention throughout this presentation.

Contact Us

Mai Mamdooh **Naira Elazab**

Sayed ELmasry **Verina Fouad** **Nermeen Reda**