

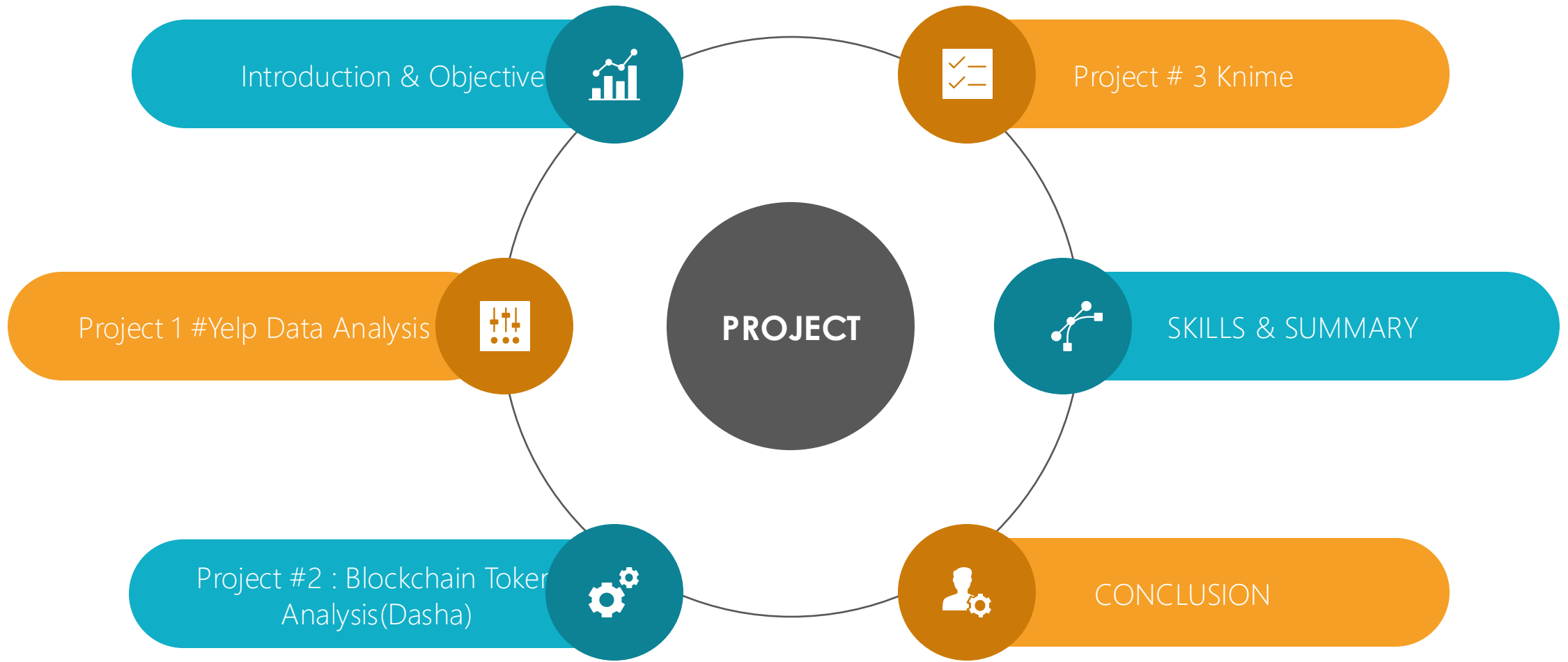


From Concept to Execution: Highlights of Our Mini Projects

Team 7

Askar Kalybay, Chu-Chun Ku, Daniel McGhee, Likhitha Kurella

AGENDA



Project 1

I want to set up a new restaurant and have yelp data.

Identify the State with the Least Proportion of Five-Star Restaurants to Open a New Restaurant in that State

- **Purpose:** A lower five-star ratio might indicate weaker competition in that state.
- **Insight:** Two states lack any five-star restaurants, but they differ in the total number of restaurants, so further analysis is needed to find the most suitable state.

state	restaurant_count	five_stars_restaurant_num	five_star_ratio
CA	44	None	None
MO	129	None	None
PA	430	8	0.02
AB	91	2	0.02
LA	130	2	0.02
ID	39	1	0.03

Calculate the Review-to-Restaurant Ratio for CA and MO to Determine Which State Has a Larger Customer Base

- **Purpose:** Use review_count as a rough indicator of customer base size to compare with the restaurant count, helping identify which state may have a higher demand.
- **Insight:** California (CA) has a much higher review-to-restaurant ratio than Missouri (MO), suggesting a larger customer base in CA, making it a better choice for a new restaurant.

state	restaurant_count	total_review_count	review_to_restaurant_ratio
CA	44	11257	255
MO	129	16824	130

city	restaurant_count	total_review_count	review_ratio
Carpinteria	6	1699	283.16666666666666
Santa Barbara	23	4316	187.652173913043
Goleta	11	1341	121.90909090909090

Identify Popular Restaurant Categories Based on Review Counts

- **Purpose:** Identify popular cuisine types by reviewing categories with high customer counts to help guide our new restaurant's menu.
- **Insight:** Pizza is widely popular, followed closely by Chinese cuisine.

categories	store_count	categories_review_count
Restaurants, Pizza	39	1313
Restaurants, Chinese	26	784
Pizza, Restaurants	24	842
Chinese, Restaurants	22	531
Mexican, Restaurants	21	1409
Restaurants, Mexican	14	1054

Project 1

Count the Number of Pizza Restaurants in Carpinteria

- **Purpose:** Assess the number of local competitors.
- **Insight:** There is only one pizza competitor, so opening a pizza restaurant here could be advantageous given the high demand.

categories	competitor_count	categories_review_count
Restaurants, French, Italian, Pizza	1	169

Review Competitor Feedback to Identify Common Pitfalls

- **Purpose:** Learn from competitor reviews to adopt best practices and avoid common issues.
- **Insight:** Feedback suggests adding televisions to appeal to a wider age range.

business_id	stars	text
ul15YlnpDsGBUZcNfm_TfQ	4	I'm glad that Michael & Lucy found a way to open up another family restaurant after years at Tony's on Linden (now closed following an ugly family dispute). The food is good and the prices are reasonable (ask about the slice and salad special). The Tuscany design is pleasant as are the owners (Lucy is always sweet and friendly). They have local beer on tap (Figueroa, Island Brew and others), 3 HD sports TVs inside, one HD sports TV outside for patio diners, a TV dedicated to surf videos and a TV dedicated to kids shows (and a classic old family friendly Ms Pac-Man game that was a hit with 2 kids when I was there). It's a tough location but I hope they make it work. Santa Claus Lane seems to be having a bit of a Renaissance after many years of being known mostly for outdated kitschy roadside attractions. Good luck!

Conclusion

- **Best State for Opening:**
 - Carpinteria , California (CA) has a larger customer base with fewer competitor indicated by a higher review-to-restaurant ratio, making it the top choice for opening a new restaurant.
- **Cuisine Focus:**
 - Pizza is the most popular, followed by Chinese food. Offering pizza could attract a large customer base.
- **Competition:**
 - Carpinteria has only one pizza restaurant, so there's minimal competition, presenting a great opportunity.
- **Differentiation Strategy:**
 - Adding televisions could help cater to a wider age demographic based on competitor feedback.

Project 2 Overview



ABOUT THE TOKEN

Dasha (VVAIFU) is a newly launched meme coin on the Solana blockchain, introduced on October 18, 2024. It has gained attention due to its rapid increase in price and trading volume, making it an attractive but risky investment.

With over 14,000 holders, Dasha presents both growth potential and risks related to market volatility and holder concentration

KEY CONCEPTS LEARNED

- Transaction Volume and Market Liquidity
- Diamond Hand
- Holder growth

INTENT AND OUTPUT OF QUERIES

The queries aimed to analyze key metrics like token price trends, transaction volume, and holder distribution for Dasha to assess its market performance.

RELEVANCE OF THE QUERIES

The queries were relevant for identifying patterns that can guide investment decisions, particularly for high-risk tokens like Dasha. By analyzing the concentration of holdings and price trends, we were able to infer the risks associated with investing in such token

INSIGHTS

Dasha's recent price fluctuations suggest that while it has significant short-term growth potential, it also carries considerable risk due to its price instability.

A large portion of tokens is held by a few addresses, which could lead to market manipulation, affecting price stability.

Project 2 - Dashboard

DASHA ANALYSIS

Persona & Goal

We are venture capitalists, and we have noticed that a meme coin - Dasha has been rising recently, with increases in both trading volume and price. Therefore, we want to evaluate its potential for appreciation and the level of risk involved.

Token Details

Token Name

Dasha

Token Name

@likhitha

18d

Ticker token details

Token details

vvaifu

Ticker

@likhitha

18d

Current Holders Token details

Token Details ..

14,331

Holders Count

@likhitha

18d

Top Holders Analysis

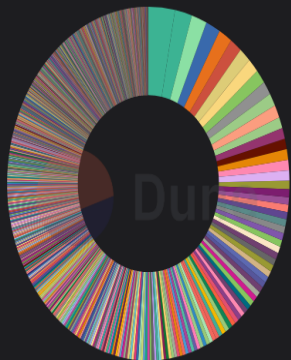
Top 1000/10 Holders

Goal: Goal: We want to see if there are any large holders and assess whether there is any risk of centralization.

Insight:

We found that there is a low centralization risk because there is no large holder.

Top 1000 holders



@likhitha

18d

Top 10 holders

holder	balance	holding
216KTuLo8kCfJam4Xmx14BtMmtn2ehYDGPfwQSVmJNdd	29815257.241959	<div></div>
8FMLCcju1FCkcC2SjHu79q5BU13qcYcwoiCAadCdYVS	24494886.254373	<div></div>
5Q544fKrFoe6tsEbD7S8EmxGTJYAKtTVhAW5Q5pge4j1	21842629.192259	<div></div>
pXgkbemDud2TCbMn6emuwzQVDKeXEqQYChRLB3zJgJh	16969427.23437	<div></div>
FSpL1gdZydJrQkngWe2S4tjBToRwpisRhChXYDQEnWPL	15953983.16368	<div></div>
7YdSXizVYVtNKUA5ihVXmUpHcFZBzT9pcoWPukUeRy7x	15285030.848405	<div></div>



10 rows Search...

@likhitha

18d

Holder Growth

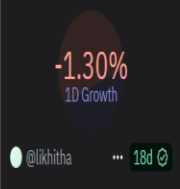
Holder Growth / Active Wallets Holding Token

Goal: Review the change in holder growth over the past 14 days & active wallets holding tokens. The charts display how the number of holders has changed, indicating the level of adoption and community expansion. "Active wallets" refers to wallets that have interacted with the token during this period, helping to track not only growth but also the engagement and activity of token holders in the market.

Insight:

We can see that the community has expanded over the past 14 days. Upon reviewing the 1-day growth rate, we believe that growth is slowing down.

1D Growth



7D Growth holder growth change

201%

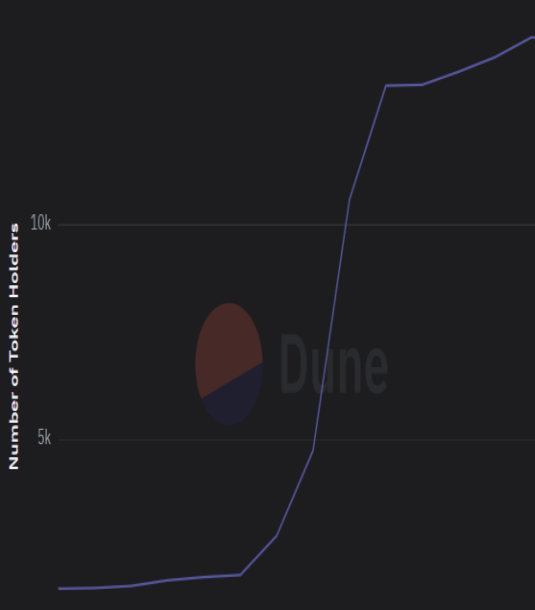
@likhitha

18d

14D Change

0.00%

Active Wallets Holding tokens Dasha-holder growth last 14 days



Project 3 – Overview



ABOUT THE PROJECT

This project focuses on data transformation and analysis using Knime. It involves extending workflows for Yelp and Memecoin (BONK) data to perform five meaningful operations that answer relevant questions, connecting results to tell a cohesive story

KEY CONCEPTS LEARNED

- Workflow design and annotation in Knime.
- Integration of multiple data transformation nodes
- Visualization in communicating insights
- Collaborative data preparation and combining datasets from multiple sources.

INTENT AND OUTPUT OF QUERIES

Intent: Identify key trends, patterns, or outliers in Yelp restaurant reviews and BONK Memecoin trades.
Output: Yelp: Filtered, grouped, and visualized insights on restaurant ratings, and made an analysis for investments.
BONK: Merged dataset with 2456 rows, transformations to explore trading patterns, identify high-volume traders, and trends on Nov 28

RELEVANCE OF THE QUERIES

- Yelp: Insights help understand customer preferences, identify top-performing restaurants, and reveal actionable trends for businesses.
- BONK: Queries provide a comprehensive view of trading activity, facilitating decisions about market trends and anomalies.

INSIGHTS

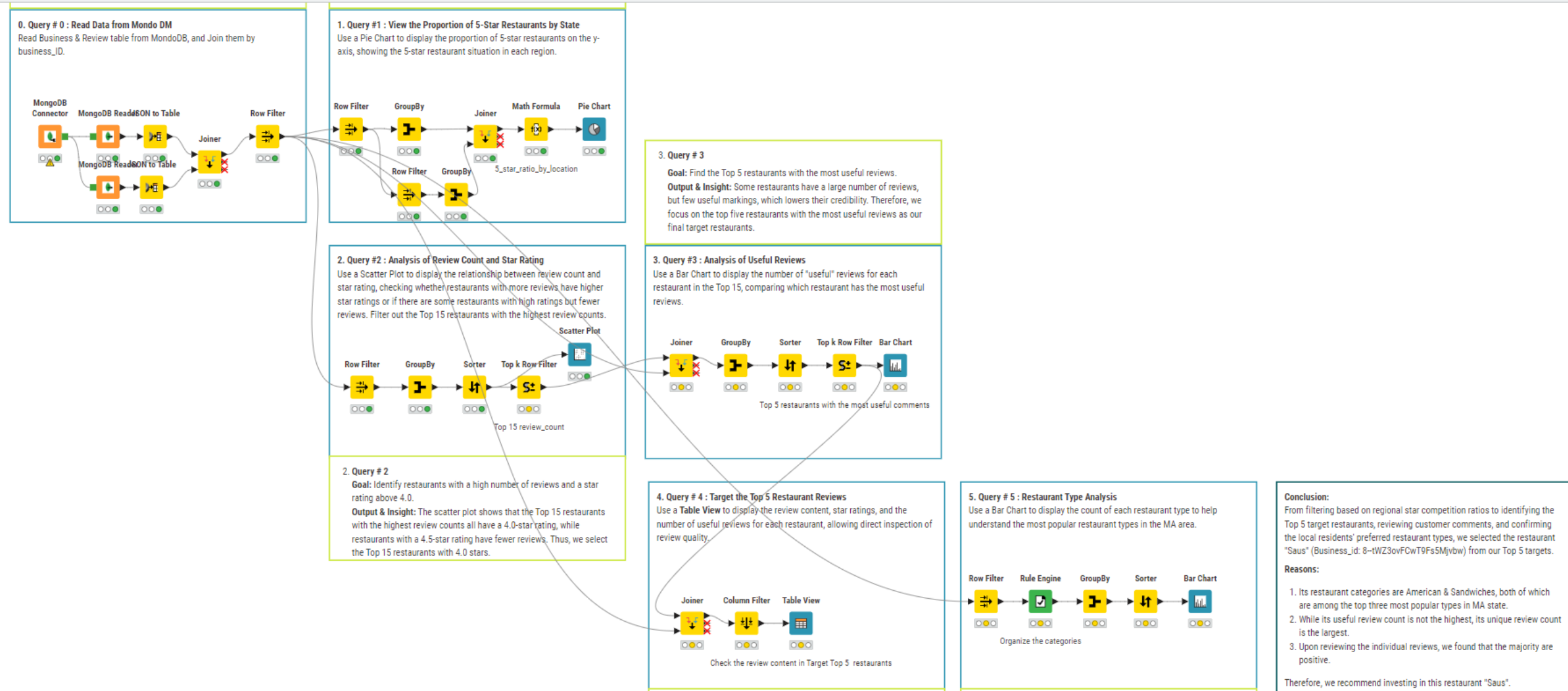
Yelp: Popular cuisines by review volume, restaurants with consistently high ratings, and regional customer preferences. It provided invaluable insights on where to invest
BONK: Key traders driving activity, patterns in trading volume throughout the day, and trends across missing attributes when combined.

Project 3 - Part 1

Persona:

We are a group of investors looking to invest in a restaurant.

We would like to refer to the star ratings and review content for evaluation.

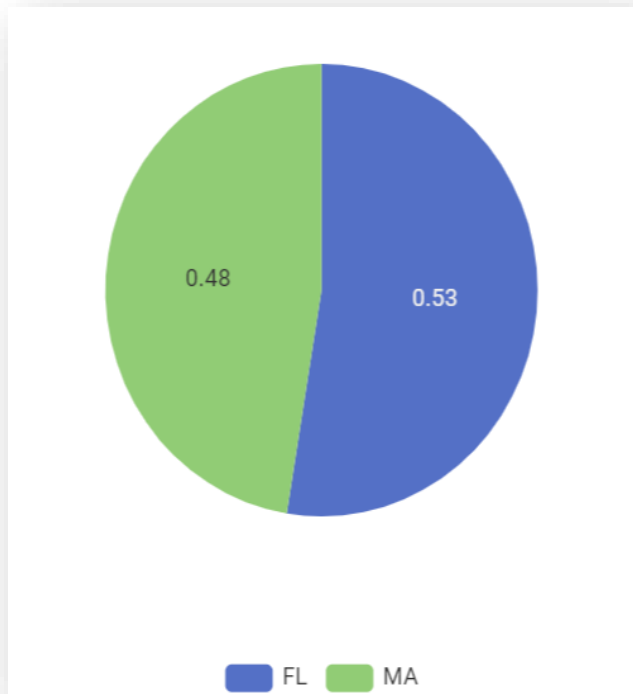


Project 3 - Part 1

Query #1

Goal: Find areas with fewer competitors by calculating the ratio of star ratings to total restaurants in the area. A lower ratio indicates fewer competitive competitors.

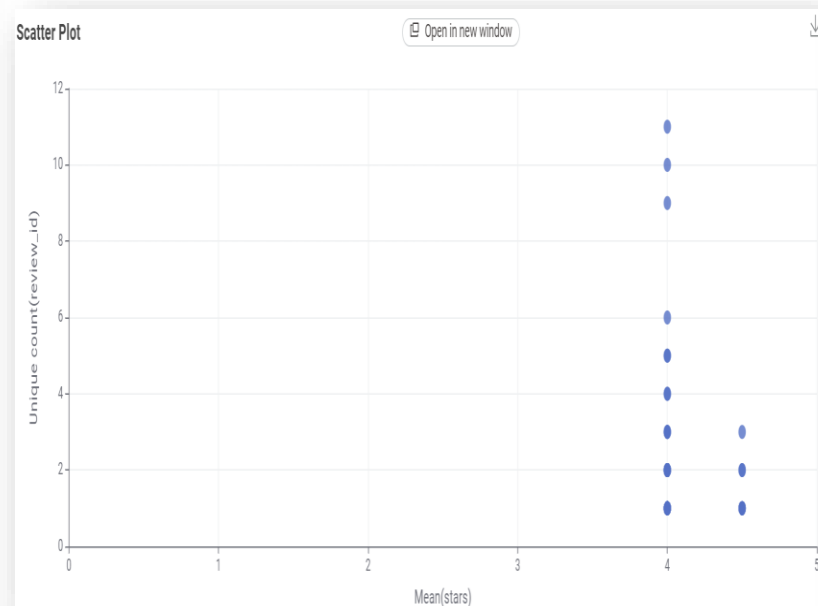
Output & Insight: The query results show that MA has fewer high-star competitors. Therefore, we will focus on setting the state to MA.



Query # 2

Goal: Identify restaurants with a high number of reviews and a star rating above 4.0.

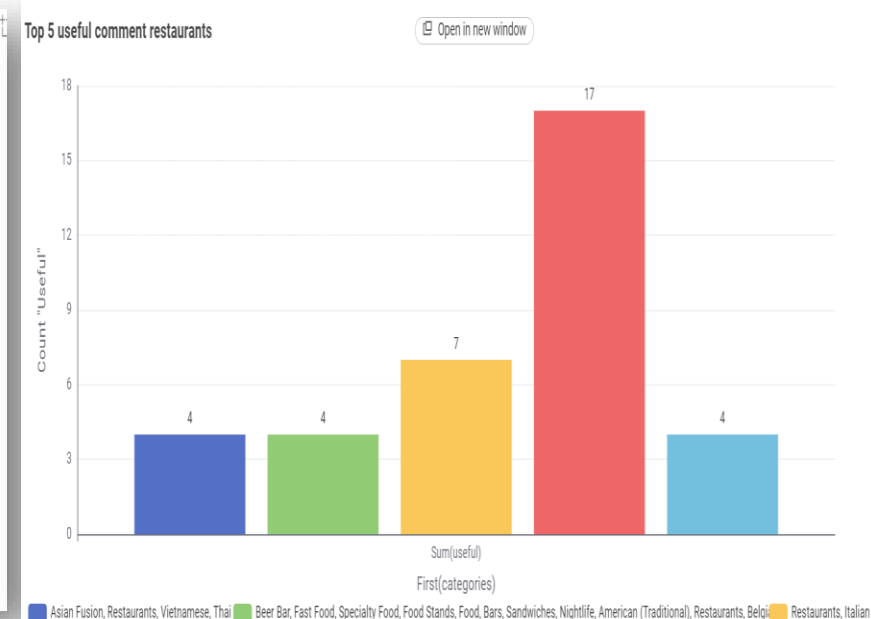
Output & Insight: The scatter plot shows that the Top 15 restaurants with the highest review counts all have a 4.0-star rating, while restaurants with a 4.5-star rating have fewer reviews. Thus, we select the Top 15 restaurants with 4.0 stars.



Query # 3

Goal: Find the Top 5 restaurants with the most useful reviews.

Output & Insight: Some restaurants have a large number of reviews, but few useful markings, which lowers their credibility. Therefore, we focus on the top five restaurants with the most useful reviews as our final target restaurants.



Project 3 - Part 1

Query # 4

Goal: Review the content of the Top 5 target restaurants to help us decide if they are suitable for investment.

Output & Insight: Some useful reviews are negative, such as "This was a really disappointing experience... I'm giving this place one star." We can filter out a few more restaurants based on this feedback.

Query # 5

Goal: Confirm the most popular restaurant types in MA state.

Output & Insight: American restaurants are the most popular in MA state. We can check if any of our Top 5 target restaurants fall into this category.

Top 5 target restaurants' comments

Rows: 26 | Columns: 6

<input type="checkbox"/>	#	RowID	b... St...	name String	S... N...	c... St...	useful Number (l...	text String
<input type="checkbox"/>	4	Row...	J1ui...	Grotto	4	Rest...	1	This was a really disappointing experience. After reading all the great reviews I thoug...
<input type="checkbox"/>	5	Row...	J1ui...	Grotto	4	Rest...	1	We went there for restaurant week and had an amazing experience. The service was a...
<input type="checkbox"/>	6	Row...	BOzd...	Beantown P...	4	Asia...	1	My favorite part of Beantown Pho is the atmosph

Conclusion:

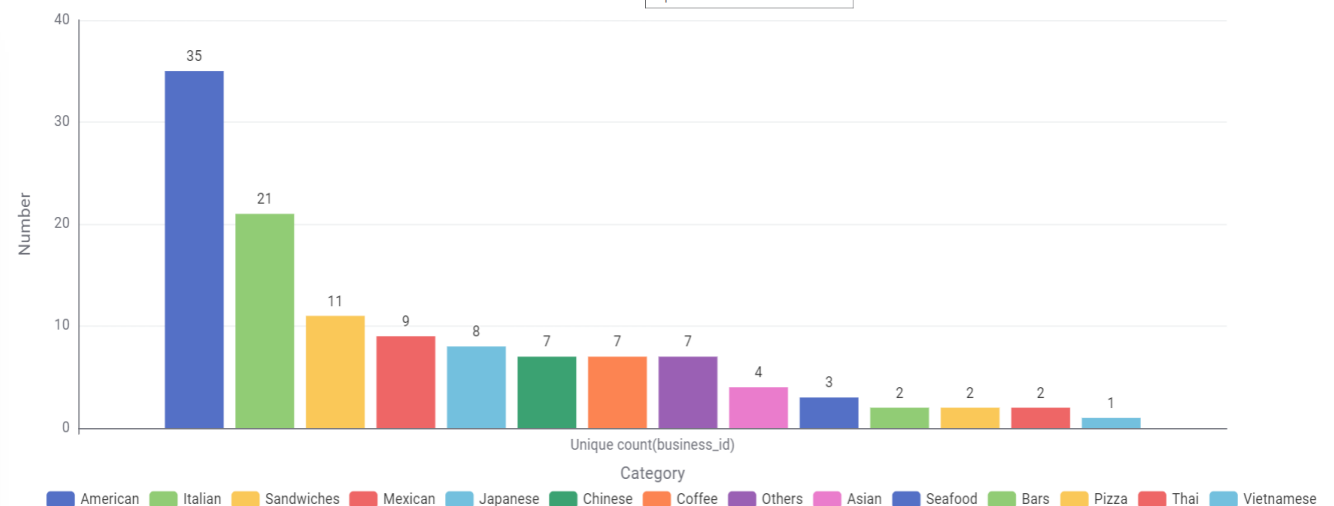
From filtering based on regional star competition ratios to identifying the Top 5 target restaurants, reviewing customer comments, and confirming the local residents' preferred restaurant types, we selected the restaurant "Saus" (Business_id: 8--tWZ3ovFCwT9Fs5Mjvbw) from our Top 5 targets.

Reasons:

1. Its restaurant categories are American & Sandwiches, both of which are among the top three most popular types in MA state.
2. While its useful review count is not the highest, its unique review count is the largest.
3. Upon reviewing the individual reviews, we found that the majority are positive.

Therefore, we recommend investing in this restaurant "Saus".

Bar Chart



Project 3 - Part 2

Overview

Objective

We were tasked to Combine BONK trade data from team members in KNIME and Perform five operations that answer a relevant question

Our Story

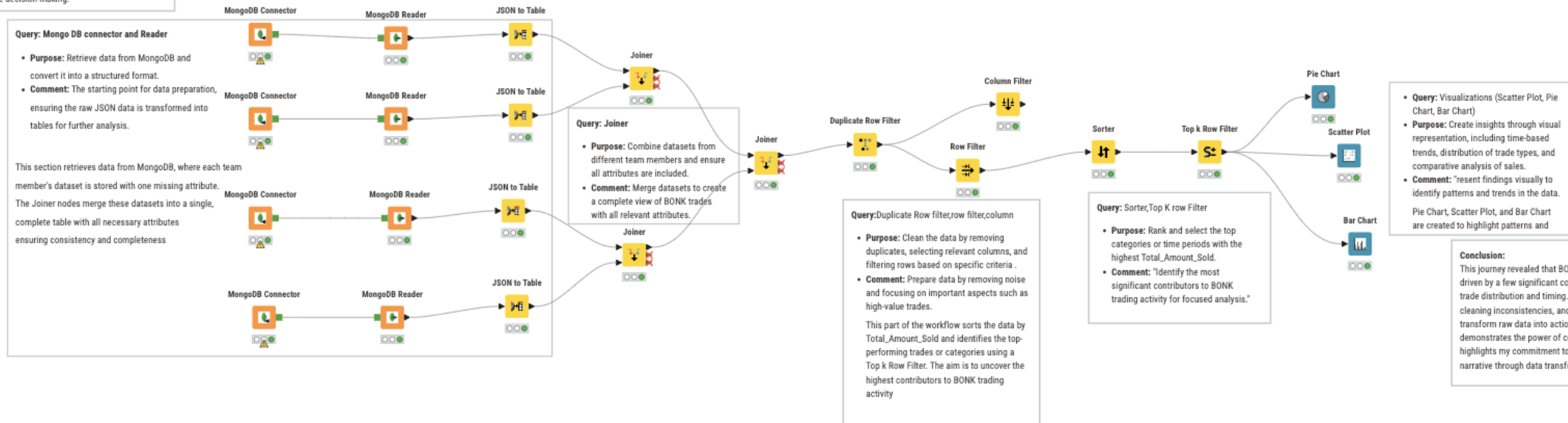
"As **an analyst** working to uncover insights from BONK trading data on November 28th, my **primary goal was to paint a clear picture of trading trends and highlight the most impactful contributors to Total_Amount_Sold**. Starting with fragmented datasets from my team, each missing a critical attribute, We took on the challenge of creating a complete and clean dataset to enable meaningful analysis. Focused on identifying top-performing trades, understanding the distribution of trading activity, and spotting trends over time. Through systematic data transformations and visualizations, I was able to tell a compelling story about the BONK trading landscape, empowering stakeholders with actionable insights for future decision-making."

Goals

Create a KNIME workflow and generate relevant insights with queries and visualizations

Project 3 - Part 2

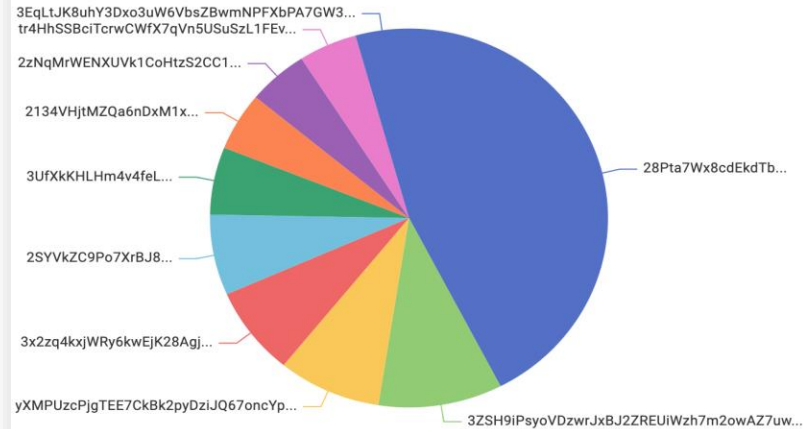
As a analyst working to uncover insights from BONK trading data on November 28th, my primary goal was to paint a clear picture of trading trends and highlight the most impactful contributors to Total_Amount_Sold. Starting with fragmented datasets from my team, each missing a critical attribute, We took on the challenge of creating a complete and clean dataset to enable meaningful analysis. Focused on identifying top-performing trades, understanding the distribution of trading activity, and spotting trends over time. Through systematic data transformations and visualizations, I was able to tell a compelling story about the BONK trading landscape, empowering stakeholders with actionable insights for future decision-making.



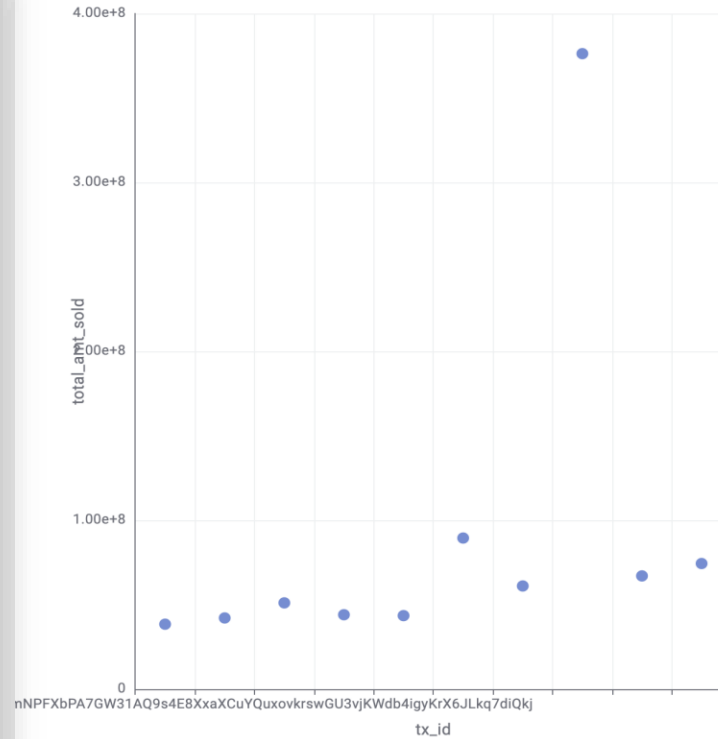
- We first combined our datasets with configured Joiner nodes
 - Then, we created queries to see specific data results
- Finally, we used those queries to visualize the relevant data

Project 3 - Part 2

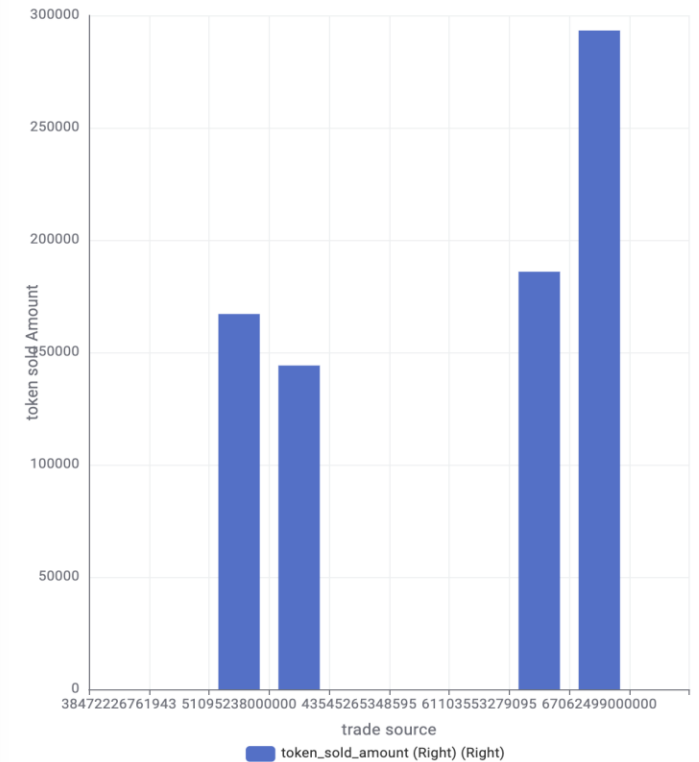
Pie Chart



Scatter Plot



Bar Chart

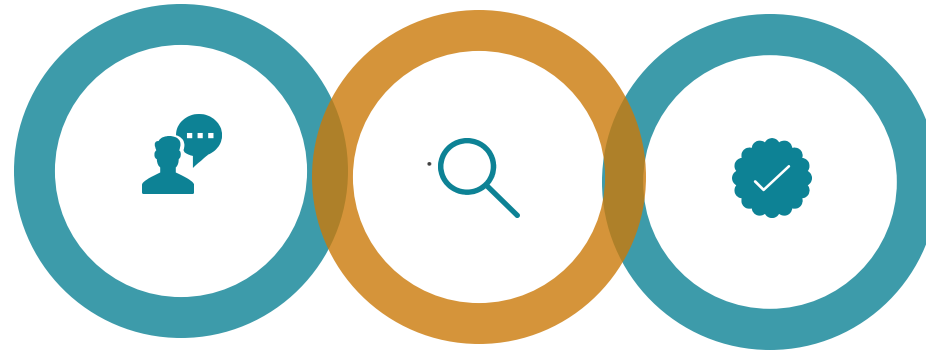


- **Explanation:** The pie chart shows the **distribution of BONK trades** across different **wallet addresses**, illustrating which wallets are contributing the most to the total trade volume.
- **Insight:** The largest segment indicates a **wallet that significantly dominates the trading activity**, while other wallets have relatively smaller shares.

- **Explanation:** The bar chart displays the **total number of tokens sold across different trade sources** (identified by transaction IDs).
- **Insight:** One trade source has the **highest token volume sold**, highlighting major contributors to the trading activity.

- **Explanation:** The scatter plot shows the **relationship between transaction IDs and the total amount of tokens sold**.
- **Insight:** There is one transaction with an **unusually high token sale amount**, suggesting a **significant trade event**.

Skills Developed



Project 1

- **Data Transformation and Integration:**
 - Transform and clean raw data (e.g., JSON) into a structured format.
- **SQL Query Writing:**
 - Develop skills in writing SQL queries to extract meaningful business insights from the Yelp dataset, helping answer practical questions related to starting a new restaurant.
- **Business Analytics and Problem Solving:**
 - Apply data analysis to solve real-world business problems, such as using Yelp data to decide where to open a new restaurant and understand customer preferences.

Project 2

- **Blockchain Analysis:**
 - Gaining insights into how meme tokens and blockchain transactions work.
- **Data Interpretation:**
 - Using tools like Dune Analytics.
- **Query Building:**
 - Developing and running efficient queries to extract relevant data on token performance-SQL

Project 3

- **KNIME Data Transformation:**
 - Proficiency in using KNIME for advanced data transformation and workflow design
- **Insight Extraction:**
 - Effective use of filtering, grouping, and sorting operations to extract meaningful insights
- **Data Merging & Integrity:**
 - Merging datasets to handle missing attributes and maintain data integrity.
- **Impactful Visualizations:**
 - Creating clear and impactful visualizations to support data-driven storytelling.

Challenges & Rewards

Challenges

- Converting complex JSON data into a structured format and porting it into an Azure SQL Server database
- Understanding financial terms while analyzing meme coins which are essential for evaluating the speculative and often unpredictable nature of meme coin markets.

Rewards

- Both Dune Analytics and KNIME provide hands-on experience with powerful tools used in the industry.
- Collaboration across team members who may be familiar with different platforms.



Thank You

24Slides