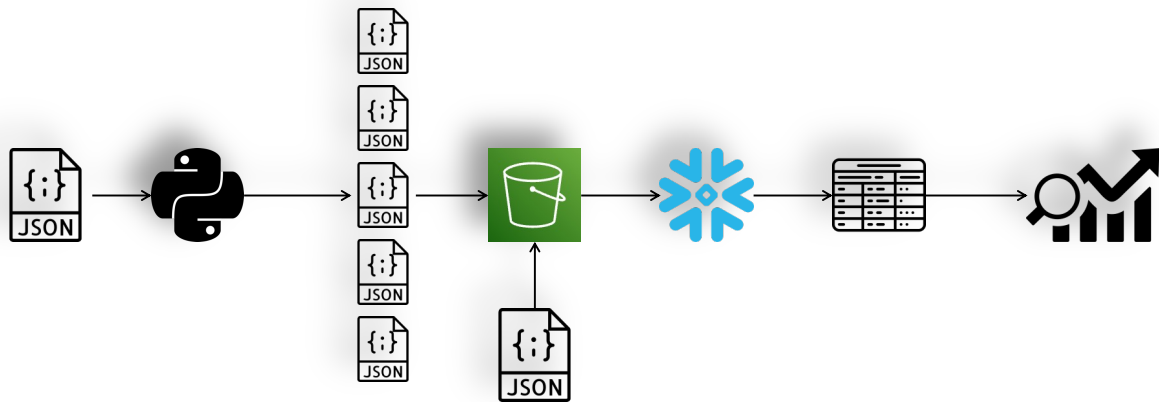



Data Analytics Pipeline Using Python, AWS and Snowflake



Steps Involved:

- **Data Download:** The data was obtained from the Yelp Business Open Dataset.
- **Data Partitioning & AWS S3 Upload:** The dataset was split into 30 parts using Python and uploaded to an AWS S3 bucket.
- **AWS S3 Setup:** The business data was added to the S3 folder, and a new AWS user with read access was created for Snowflake integration.
- **Data Fetch & Transformation in Snowflake:** Data in JSON format was loaded into Snowflake using the VARIANT datatype, and two tables (reviews and businesses) were created from JSON keys.
- **Sentiment Analysis Using UDF:** A User Defined Function (UDF) in Snowflake was created to analyze sentiment of review text using Python's TextBlob, categorizing sentiments as 'Positive', 'Negative', or 'Neutral'.
- **Business Question Answers:** After table creation, 10 business-related questions were answered using SQL queries in Snowflake.

• YELP Dataset

 data licensing

Products ▾

Solutions ▾

Resources ▾

Book a demo

Download

This download contains 1 compressed TAR file (4.35 GB).
Uncompressed it contains 1 JSON file, 1 PDF and 5 JSON files (8.65 GB).
Documentation is included.

Download JSON



This download contains one compressed TAR file (7.45 GB).
Uncompressed it contains 1 JSON file, 1 text file, 1 PDF and
1 folder containing 200,000 photos (7.11 GB).

Download photos

• Python Data Division

```
[12] Generate + Code + Markdown | Run All | Clear All Outputs | Outline | Select Kernel Python

[12] num_of_files = 30
lines_per_file = total_lines // num_of_files

[13] Python

[13] print(f"Total lines: {total_lines}")
print(f"Lines per file: {lines_per_file}")

[14] Python

... Total lines: 6990200
Lines per file: 233009

[14] # now split the file into smaller files
with open("yelp_dataset/yelp_academic_dataset_review.json", "r", encoding="utf8") as f:
    for i in range(num_of_files):
        with open(f"yelp_dataset/yelp_academic_dataset_review_{i+1}.json", "w", encoding="utf8") as out_file:
            for j in range(lines_per_file):
                line = f.readline()
                if not line:
                    break
                out_file.write(line)
            print(f"File {i+1} done")
print("All files done")

[15] Python

... File 1 done
File 2 done
File 3 done
File 4 done
File 5 done
File 6 done
File 7 done
File 8 done
File 9 done
```

• AWS Data Upload

🟢 Upload succeeded
For more information, see the Files and folders table.


Files and folders

Configuration

Files and folders (30 total, 5.0 GB)

Find by name							< 1 2 3 >	
Name	Folder	Type	Size	Status	Error			
yelp_academic_dataset_review...	-	application/json	176.5 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	167.5 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	164.0 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	176.6 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	167.7 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	164.4 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	176.2 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	166.6 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	163.9 MB	🟢 Succeeded	-			
yelp_academic_dataset_review...	-	application/json	177.5 MB	🟢 Succeeded	-			

• Business data fetched from AWS



Load sample data with SQ...

2025-04-25 7:42pm

2025-04-26 10:59am

2025-04-26 11:16am

+

-

Databases

Worksheets

Search objects

LEARNSQL

SNOWFLAKE

SNOWFLAKE_SAMPLE_DATA

Create

Home

Search

Projects

Worksheets

Notebooks

Streamlit

Dashboards

App Packages

Data

Data Products

AI & ML

Monitoring

Admin

\$398 credits left

Trial ends in 29 days

AP Aditya Pandey

ACCOUNTADMIN

LEARNSQL_PUBLIC

Settings

Code Versions

```
1 create or replace table yelp_bussiness(bussiness_text variant)
2
3 Copy into yelp_bussiness
4 from 's3://namastesql-aditya/yelp/yelp_academic_dataset_business.json'
5 credentials = (
6   AWS_KEY_ID = '
7   AWS_SECRET_KEY = '
8 )
9 FILE_FORMAT = (TYPE = JSON);
```

Results

Chart

file	status	# rows_parsed	# rows_loaded	# error
s3://namastesql-aditya	LOADED	150346	150346	

Query Details

Query duration 8.7s

Rows 1

Query ID 01bbf1ec-0306-5af2-00...

Show more

file

100% filled

Ask Copilot

- Review data fetched from AWS

The screenshot shows the Snowflake web interface. On the left is a navigation sidebar with options like Home, Search, Projects, Worksheets, Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes. The left pane shows a database hierarchy: LEARNSQL > SNOWFLAKE > SNOWFLAKE_SAMPLE_DATA. The middle pane contains a SQL query:

```
1 create or replace table yelp_reviews (review_text variant)
2
3 Copy into yelp_reviews
4 from 's3://namastesql-aditya/yelp/'
5 credentials = (
6   AWS_KEY_ID = 'REDACTED'
7   AWS_SECRET_KEY = 'REDACTED'
8 )
9 FILE_FORMAT = (TYPE = JSON);
```

 The right pane shows the query results in a table with columns: file, status, # rows_parsed, # rows, and Query Details. The table lists 10 rows, all with a status of 'LOADED' and 233009 rows parsed. The Query Details pane on the right shows a query duration of 35s and a query ID of 01bbf1e4-0306-58cd-0...

file	status	# rows_parsed	# rows	Query Details
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	Query duration 35s
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	Rows 30
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	Query ID 01bbf1e4-0306-58cd-0...
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	Show more
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	file 100% filled
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	status Ask Copilot
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	
s3://namastesql-aditya/yelp/yelp_academic_dataset_review	LOADED	233009	233009	

- UDF

The screenshot shows the Snowflake web interface. The left sidebar is the same as in the previous screenshot. The middle pane shows a database hierarchy: LEARNSQL > INFORMATION_SCHEMA > PUBLIC > Tables > REVIEW. The right pane contains a SQL query to create a UDF:

```
1 create or replace function analyse_sentiment(text string)
2 returns string
3 language python
4 runtime_version = '3.8'
5 packages = ('textblob')
6 handler = 'sentiment_analyzer'
7 as $$
8 from textblob import TextBlob
9 def sentiment_analyzer(text):
10   analysis = TextBlob(text)
11   if analysis.sentiment.polarity > 0.3:
12     return 'Positive'
13   elif analysis.sentiment.polarity < 0.3 and analysis.sentiment.polarity > -0.3:
14     return 'Neutral'
15   else:
16     return 'Negative'
17
18 $$
```

 The bottom pane shows the results of the UDF in a table with columns: REVIEW and ANALYSE_SENTIMENT(REVIEW). The table lists 6 rows of reviews and their corresponding sentiment analysis results. The Query Details pane on the right shows a query duration of 1.7s and a query ID of 01bbf2a6-0306-5b21-0...

REVIEW	ANALYSE_SENTIMENT(REVIEW)
I'm not sure how I feel about this.	Neutral
I hated every minute of it.	Negative
Incredible performance by the cast!	Positive
Absolutely fantastic service!	Positive
I have no strong opinion on this.	Neutral
The product broke after one use.	Neutral

• JSON to structured table (Review Table)

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes Home, Search, Projects, Worksheets (selected), Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes. The left pane shows a list of worksheets, with '2025-04-26 1:34pm' selected. The middle pane displays a SQL query in the 'LEARNSQL.PUBLIC' database:

```
1 create or replace table tbl_yelp_reviews as
2 select
3   review_text:business_id:: string as business_id,
4   review_text:date:: date as review_date,
5   review_text:stars:: number as stars,
6   review_text:user_id:: string as user_id,
7   review_text:text:: string as review_text,
8   analyse_sentiment(review_text) as sentiment
9 from yelp_reviews
10
11
12 | select * from tbl_yelp_reviews limit 10
```

The right pane shows the query results in a table format:

	BUSINESS_ID	REVIEW_DATE	STARS	USER_ID
1	xAwES1jxVnoLmIPGRVu30Q	2017-08-26	4	pl6mQmn49cCyZZakIKqLaw
2	OUTw115MSuk-MoB_QEmrog	2018-12-20	4	4_H7deukZvhrdEi@zabgOw
3	M6ap9LXMEZmHf6NhhNpDQA	2019-01-12	4	ALCxAt7aKE5C6R6arZyig
4	Vw2pDOOU2AMHL_LD_U6A	2022-01-19	5	zwaCKuPLsWM4Cq5K0yUQQ
5	z2VVTg2OaEYmxN4n6RHMJA	2019-02-18	5	Mhj_HpX4QISbSaPZURcgkw
6	XvCGAQLC9y6lfHu1R_L3g	2016-03-20	5	lrl3z3C-6l6a-O5xfKm1Fg

Below the table, the 'Query Details' pane shows: Query duration: 229ms, Rows: 10, Query ID: 01bbf46c-0308-58cd-0... An 'Ask Copilot' button is visible at the bottom right.

• JSON to structured table (Business Table)

The screenshot shows the Snowflake web interface. The left navigation menu is the same as in the previous screenshot. The main workspace shows the 'LEARNSQL.PUBLIC' database with a SQL query:

```
1 create table tbl_yelp_business as
2 select
3   business_text:business_id:: string as business_id,
4   business_text:city:: string as city,
5   business_text:state:: string as state,
6   business_text:review_count:: number as review_count,
7   business_text:categories:: string as categories,
8   business_text:stars:: number as stars
9 from yelp_business |
```

The right pane shows the query results in a table format:

	status
1	Table TBL_YELP_BUSINESS successfully created.

The 'Query Details' pane shows: Query duration: 938ms, Rows: 1, Query ID: 01bbf3b2-0308-5485-0... An 'Ask Copilot' button is visible at the bottom right.

1. Find the Number of Businesses in Each Category

The screenshot shows the Snowflake web interface. On the left is a sidebar with navigation options like Home, Search, Projects, Worksheets, Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes. The left pane shows a database schema with tables like ORDERS, REVIEW, TBL_YELP_BUSSINESS, TBL_YELP_REVIEWS, YELP_BUSSINESS, and YELP_REVIEWS. The middle pane contains a SQL query:

```
1 /* 1. Find Number business in each category*/
2 with cte as (
3 select business_id, trim(A.value) as category from tbl_yelp_bussiness, lateral
4 split_to_table(categories, ',' ) A)
5 select category, count(*) count from cte group by 1 order by 2 desc
6 --select * from tbl_yelp_bussiness
```

 The right pane displays the query results in a table with two columns: CATEGORY and # COUNT. The results show 7 categories: Restaurants (52268), Food (27781), Shopping (24395), Home Services (14356), Beauty & Spas (14292), Nightlife (12281), and Health & Medical (11890). A 'Query Details' panel on the far right shows a query duration of 305ms and 1.3K rows.

CATEGORY	# COUNT
Restaurants	52268
Food	27781
Shopping	24395
Home Services	14356
Beauty & Spas	14292
Nightlife	12281
Health & Medical	11890

2. Find Top 10 Users Who Have Reviewed the Most Businesses in "Restaurant" Category

The screenshot shows the Snowflake web interface with a different SQL query. The query is:

```
1 /*2. Find top 10 users who have reviewed the most businesses in "Restaurant" Category*/
2 select r.user_id, count(distinct r.business_id) number_of_review from tbl_yelp_reviews r join
3 tbl_yelp_bussiness b on b.business_id=r.business_id
4 where b.categories like '%Restaurant%'
5 group by 1
6 order by 2 desc
7 limit 10
```

 The results table shows the top 10 users by the number of reviews. The columns are USER_ID and # NUMBER_OF_REVIEW. The top user is -G7Zk1wIWBmDOKRy_sCw with 1202 reviews. Other users include _BcWyKQL16ndpBdggh2kNA (1166), fr1Hz2acAb30al3i6dyKNg (1058), 1HM81n6n4iPIFUsd2Lokhw (1009), Xw7JgGf0WNVN16s_5KZfA (926), ET8n-v7gWYqZhuR6OcdNw (891), and pou3BbkslozfH50rxmnMew (849). The 'Query Details' panel shows a query duration of 1.1s and 10 rows.

USER_ID	# NUMBER_OF_REVIEW
-G7Zk1wIWBmDOKRy_sCw	1202
_BcWyKQL16ndpBdggh2kNA	1166
fr1Hz2acAb30al3i6dyKNg	1058
1HM81n6n4iPIFUsd2Lokhw	1009
Xw7JgGf0WNVN16s_5KZfA	926
ET8n-v7gWYqZhuR6OcdNw	891
pou3BbkslozfH50rxmnMew	849

3. Find Most Popular Categories of Business (Based on Number of Reviews)

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes 'Create', 'Home', 'Search', 'Projects', 'Worksheets' (selected), 'Notebooks', 'Streamlit', 'Dashboards', 'App Packages', 'Data', 'Data Products', 'AI & ML', 'Monitoring', and 'Admin'. A status bar at the bottom left shows '\$392 credits left' and 'Trial ends in 29 days'. The main workspace is titled '2025-04-26 6:54pm'. The 'Databases' tab is active, showing a tree view of the 'PUBLIC' schema with tables: 'ORDERS', 'REVIEW', 'TBL_YELP_BUSINESS', 'TBL_YELP_REVIEWS' (selected), 'YELP_BUSINESS', and 'YELP_REVIEWS'. The 'Snowflake Sample Data' section shows 'TBL_YELP_REVIEWS' with 7.0M rows. The query editor on the right contains the following SQL:

```
1 /*3. Find top 3 most recent reviews for each business*/
2 select B.business_name,review_date,review_text,sentiment
3 from tbl_yelp_reviews R join tbl_yelp_business B on B.BUSINESS_ID=R.business_id
4 qualify row_number() over(partition by b.business_id order by review_date desc) <=3
5
6
```

The 'Results' tab shows a table with 7 rows and 3 columns: BUSINESS_NAME, REVIEW_DATE, and REVIEW_TEXT. The query details on the right show a duration of 7.4s and 451.0K rows. An 'Ask Copilot' button is visible.

BUSINESS_NAME	REVIEW_DATE	REVIEW_TEXT
Jimmy The Greek's	2018-12-19	If only we found this place sooner. My wife and I were l
Jimmy The Greek's	2018-10-29	Absolutely awful. Our Spanikopita appetizer was burne
Jimmy The Greek's	2018-06-09	My husband had the 6 oz pepperloin meal and was ver
Spectrum	2022-01-06	This revele is for the Tarpon Springs location on US 19
Spectrum	2022-01-03	Spectrum strikes again! This horrible company sent me
Spectrum	2021-12-23	This is for the Clearwater location at Countryside Westf
Dollar Tree Store	2021-05-24	Need more than one person checking out very long line

4. Find Top 3 Most Recent Reviews for Each Business

The screenshot shows the Snowflake web interface. On the left, the navigation menu is the same as in the previous screenshot. The main workspace is titled '2025-04-26 6:54pm'. The 'Databases' tab is active, showing a tree view of the 'LEARNSQL' schema with tables: 'INFORMATION_SCHEMA', 'PUBLIC', and 'Tables' (selected). The 'Snowflake Sample Data' section shows 'TBL_YELP_REVIEWS' with 7.0M rows. The query editor on the right contains the following SQL:

```
1 /*4. Find top 3 most recent reviews for each business*/
2 select B.business_name,review_date,review_text,sentiment
3 from tbl_yelp_reviews R join tbl_yelp_business B on B.BUSINESS_ID=R.business_id
4 qualify row_number() over(partition by b.business_id order by review_date desc) <=3
5
6
```

The 'Results' tab shows a table with 7 rows and 3 columns: BUSINESS_NAME, REVIEW_DATE, and REVIEW_TEXT. The query details on the right show a duration of 7.4s and 451.0K rows. An 'Ask Copilot' button is visible.

BUSINESS_NAME	REVIEW_DATE	REVIEW_TEXT
Jimmy The Greek's	2018-12-19	If only we found this place sooner. My wife and I were l
Jimmy The Greek's	2018-10-29	Absolutely awful. Our Spanikopita appetizer was burne
Jimmy The Greek's	2018-06-09	My husband had the 6 oz pepperloin meal and was ver
Spectrum	2022-01-06	This revele is for the Tarpon Springs location on US 19
Spectrum	2022-01-03	Spectrum strikes again! This horrible company sent me
Spectrum	2021-12-23	This is for the Clearwater location at Countryside Westf
Dollar Tree Store	2021-05-24	Need more than one person checking out very long line

5. Find the Month with the Highest Number of Reviews

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes Home, Search, Projects, Worksheets (selected), Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes. The left pane shows the database schema with 'PUBLIC' as the selected database, containing tables like 'ORDERS', 'REVIEW', 'TBL_YELP_BUSINESS', 'TBL_YELP_REVIEWS', 'YELP_BUSINESS', and 'YELP_REVIEWS'. The middle pane contains a SQL query:

```
1 /*5. Find percentage of 5 star reviews for each business*/
2 select
3   B.BUSINESS_ID,
4   B.BUSINESS_NAME,
5   sum(case when star = 5 then 1 else 0 end)*100.0/count(star) percentage_of_5_star
6 from tbl_yelp_reviews R join tbl_yelp_business B
7 on R.BUSINESS_ID=B.BUSINESS_ID
8 group by ALL
9
```

 The right pane displays the query results in a table with columns: BUSINESS_ID, BUSINESS_NAME, and PERCENTAGE_OF_5_STAR. The results show data for various businesses, with the highest percentage of 5-star reviews being 56.597222 for 'Bill Pearce Courtesy Honda'. The bottom right pane shows query details: Query duration 23ms, Rows 150.3K, and Query ID 01bbf428-0306-5b95-0...

BUSINESS_ID	BUSINESS_NAME	PERCENTAGE_OF_5_STAR
59QYxLqYMMAnAxxk3tg	Bill Pearce Courtesy Honda	56.597222
-xJAOOYVnUjHxNcpUmQ	Yellowstone Clothing Inc	26.086957
zwngCMuZyFX46mL3pIdYjg	Cup Cafe	51.167964
tW8-5eqjnbSWoydRV_0Pg	Doggie Style Pets	50.000000
2X0fTGP6cdezBN5WC263A	My Little Secret Waxing Boutique	74.509804
URURwZcedT6sAXbeTuxgOA	Bait	60.995851
vyu-Ur33_Tk8Km2ocghFCg	Crown Eurocars	30.769231

6. Find Percentage of 5-Star Reviews for Each Business

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes Home, Search, Projects, Worksheets (selected), Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes. The left pane shows the database schema with 'PUBLIC' as the selected database, containing tables like 'ORDERS', 'REVIEW', 'TBL_YELP_BUSINESS', 'TBL_YELP_REVIEWS', 'YELP_BUSINESS', and 'YELP_REVIEWS'. The middle pane contains a SQL query:

```
1 /*6. Find percentage of 5 star reviews for each business*/
2 select
3   B.BUSINESS_ID,
4   B.BUSINESS_NAME,
5   sum(case when star = 5 then 1 else 0 end)*100.0/count(star) percentage_of_5_star
6 from tbl_yelp_reviews R join tbl_yelp_business B
7 on R.BUSINESS_ID=B.BUSINESS_ID
8 group by ALL
9
```

 The right pane displays the query results in a table with columns: BUSINESS_ID, BUSINESS_NAME, and PERCENTAGE_OF_5_STAR. The results show data for various businesses, with the highest percentage of 5-star reviews being 56.597222 for 'Bill Pearce Courtesy Honda'. The bottom right pane shows query details: Query duration 23ms, Rows 150.3K, and Query ID 01bbf428-0306-5b95-0...

BUSINESS_ID	BUSINESS_NAME	PERCENTAGE_OF_5_STAR
59QYxLqYMMAnAxxk3tg	Bill Pearce Courtesy Honda	56.597222
-xJAOOYVnUjHxNcpUmQ	Yellowstone Clothing Inc	26.086957
zwngCMuZyFX46mL3pIdYjg	Cup Cafe	51.167964
tW8-5eqjnbSWoydRV_0Pg	Doggie Style Pets	50.000000
2X0fTGP6cdezBN5WC263A	My Little Secret Waxing Boutique	74.509804
URURwZcedT6sAXbeTuxgOA	Bait	60.995851
vyu-Ur33_Tk8Km2ocghFCg	Crown Eurocars	30.769231

7. Find the Top 5 Most Reviewed Businesses in Each City

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes Home, Search, Projects, Worksheets (selected), Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main panel displays a query in the LEARNSQL_PUBLIC database. The query is as follows:

```
1 /*7. Find the top 5 most reviews businesses in each city*/
2 select
3     b.business_id,
4     b.business_name,
5     b.city,
6     count(*) review_count_
7 from tbl_yelp_business B
8 join tbl_yelp_reviews R on R.business_id=b.business_id
9 group by all
10 qualify row_number() over(partition by city order by review_count_ desc) <=5
11 order by 3
```

The results table shows the top 5 most reviewed businesses in each city. The columns are BUSINESS_ID, BUSINESS_NAME, CITY, and REVIEW_COUNT. The results are as follows:

BUSINESS_ID	BUSINESS_NAME	CITY	REVIEW_COUNT
545	MWAZAdxag2i64jwv57g	Chester	895
546	6wp8GLsJaaox7ab4wrQXLA	Chester	133
547	Jw-ENQw-OGGwpM3nA3yww	Chester	53
548	lXAXmy9neZ5_UgmrKcJaa	Chester	101
549	D0oog_GL2Rzc5Hh2QLP3A	Chester	182
550	zekGLwMoPhREmlwtswX-w	Chester Heights	7
551	d4H6t7L1LQOJS5L4IJQ	Chester Heights	20

The right sidebar shows the Query Details, including Query duration (577ms), Rows (4.2K), and Query ID (01bbf43b-0306-5b95-0-...). There is also an Ask Copilot button.

8. Find Average Rating of Businesses that Have At Least 100 Reviews

The screenshot shows the Snowflake web interface. On the left, the navigation menu includes Home, Search, Projects, Worksheets (selected), Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main panel displays a query in the LEARNSQL_PUBLIC database. The query is as follows:

```
1 /* 8. Find average rating of business that have at least 100 reviews*/
2 select b.business_name, avg(r.star)
3 from tbl_yelp_business B
4 join tbl_yelp_reviews R on R.business_id=b.business_id
5 group by 1
6 having count(*)>=100
7 order by 1
```

The results table shows the average rating of businesses that have at least 100 reviews. The columns are BUSINESS_NAME and AVG(R.STAR). The results are as follows:

BUSINESS_NAME	AVG(R.STAR)
320 Alberto's Cheese & Wine Bistro	4.647059
321 Albertsons	2.953052
322 Albertsons Market Street	4.305085
323 Albita's Mexican Restaurant	4.522059
324 Alcazar Tapas Bar	4.206452
325 Alden Suites A Beachfront Resort	3.666667
326 Aldertos Fresh Mexican Food	3.633466

The right sidebar shows the Query Details, including Query duration (585ms), Rows (14.6K), and Query ID (01bbf442-0306-5b95-0-...). There is also an Ask Copilot button.

9. List Top 10 Users Who Have Written the Most Reviews Along with the Businesses They Reviewed

The screenshot shows the Snowflake SQL interface. The left sidebar contains navigation options like Home, Search, Projects, Worksheets, Notebooks, Streamlit, Dashboards, App Packages, Data, Data Products, AI & ML, Monitoring, and Admin. The main workspace is divided into three panes: a search pane on the left showing a tree of databases and tables, a central SQL editor, and a results pane on the right.

SQL Query:

```
1  /* 9. List top 10 users who have written the most reviews along with business they reviewed */
2  with top_10_reviewers as (
3    select user_id, count(*) no_of_review
4    from tbl_yelp_reviews
5    group by 1
6    order by 2 desc
7    limit 10)
8  select B.Business_name, R.user_id from top_10_reviewers TR
9  join tbl_yelp_reviews R on r.user_id=TR.user_id
10 join tbl_yelp_business B on B.BUSINESS_ID=R.BUSINESS_ID
11 order by 2,1
```

Results Table:

	BUSINESS_NAME	USER_ID
1	1911 Smoke House Barbeque	-G7Zk1wiWBmDOKRy.sCw
2	1st NE International Market	-G7Zk1wiWBmDOKRy.sCw
3	1st NE International Market	-G7Zk1wiWBmDOKRy.sCw
4	1st Oriental Supermarket	-G7Zk1wiWBmDOKRy.sCw
5	21 Chinese Kitchen	-G7Zk1wiWBmDOKRy.sCw
6	232 Corner Cafe	-G7Zk1wiWBmDOKRy.sCw
7	232 Corner Cafe	-G7Zk1wiWBmDOKRy.sCw

Query Details:

- Query duration: 788ms
- Rows: 173K
- Query ID: 01bbf44b-0306-5b21-0...

10. Find Top 10 Businesses with the Highest Positive Sentiment Reviews

The screenshot shows the Snowflake SQL interface with a query to find the top 10 businesses by positive sentiment. The layout is similar to the previous screenshot, with a search pane, a SQL editor, and a results pane.

SQL Query:

```
1  /*10. Find top 10 businesses with highest with highest positive sentiment reviews*/
2  select B.BUSINESS_ID, B.BUSINESS_NAME, count(*) count_positive_sentiment from tbl_yelp_business B
3  join tbl_yelp_reviews R on R.BUSINESS_ID=B.BUSINESS_ID
4  where R.sentiment = 'Positive'
5  group by 1,2
6  order by 3 desc
7  limit 10
```

Results Table:

	BUSINESS_ID	BUSINESS_NAME	## COUNT_POSITIVE_SENTIMENT
1	ac1AeYgs8Z4_e2X5M3if2A	Oceana Grill	4636
2	_ab5OqdWokODd86XorBitw	Acme Oyster House	3422
3	_C7QIQQc47AOEv4PE3Kong	Commander's Palace	2845
4	ytynqUub3hjKeJfRj5Tshw	Reading Terminal Market	2817
5	cBNrLz4EDhiscSibOI8uAw	Ruby Slipper - New Orleans	2713
6	VQcCL9PINL_wkGf-uf3fjg	Royal House	2647
7	GBTPC53ZG1ZBY3DT8Mbcw	Luke	2587

Query Details:

- Query duration: 668ms
- Rows: 10
- Query ID: 01bbf451-0306-5485-0...