A close-up, low-angle shot of a person's thumb swiping left on a smartphone screen. The screen is brightly lit, casting a glow on the person's thumb and the surrounding dark environment. In the background, several other smartphones are visible, their screens glowing with various shades of green and blue, creating a bokeh effect.

SWIPE RIGHT ON DATA: REIMAGINING DATING APPS FOR THE FUTURE

DASHIA PENNINGTON

WHY DATING APP REVIEWS MATTER

Problem Statement: What problems are users consistently experiencing across dating apps, and what underlying issues do these reviews reveal?

- Dating apps **generate millions of users interactions and experiences every day.**
- Users often express frustrations around **matching quality, algorithm inconsistency, safety concerns, and bot/catfishing profiles.**
- Reviews expose deeper issues with **trust, governance, and platform accountability.**
- Understanding these patterns allows platforms to improve user protection, safety, and retention.

Objective:

- Identify dominant user frustration across Tinder, Bumble, and Hinge.
- Extract insights using AWS-based data engineering tools.

Key Insight: *The core issues aren't romantic instead they're structural. User frustration is a governance, safety, and data management issue disguised as "dating problems."*



KEY STAKEHOLDERS IN TRUST, SAFETY & GOVERNANCE

- **Primary:** Users (safety, emotional trust, harassment prevention, psychological well-being).
- **Secondary:** Dating platforms (retention, liability, accountability and moderate usage).
- **Regulatory:** General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) data protection compliance.
- **Academic Research/Public:** Policy professionals, psychological and behavioral insights.

Beware of scammers

★★★★★ Mon · cytisthebest

Unfortunately, while I've met a few nice people and in general enjoy this site, there are way too many scammers. It's discouraging because anyone with half a brain can pick them out.

It's alright

★★★★★ Mon · Stevie McKeef

I live in a large city with over 3 million people. Somehow it says there are no people around me. I have no filters set so I don't know where everyone is hiding.

Money hungry!

★★★★★ Mon · Zjnae

The people who made this app make it hard to find matches unless you pay but when you go to pay it's so expansive! \$70 for one month 30 for one week. Don't get it.

Kicks you off for no reason

★★★★★ Mon · Sarah4557

Worst app for dating. Kicked me off for no reason right after I signed up and paid. Has happened twice. Don't sign up happens to many people look at reviews

Be aware of scams

★★★★★ Mon · AppReviewer2270

6 years ago this app was awesome and safe. Now its full of scammers (girls), be aware guys. Once conversation starts , immediately they ask u for a snapchat, where they u know, do stuff and try to extort u or ask for money lol. Don't fall for that. This message is for US based app users. In Europe its much different. Better. Cant believe Tinder cant get hold of this, and make application safe to use. Want my money back.

Wow

★★★★★ Mon · Go no,!?,?

The girl I found was lesbian... im a guy

No one is real

★★★★★ Mon · Mister valerius

99% of the profiles on here are fake

So many bots...

★★★★★ Mon · mnrockclimber

Every third profile is a bot. They don't even try to hide it. They all follow the same basic formula so I don't know why Tinder doesn't weed them out. This is the worst dating app of them all.

Truly a sadistic and inhumane app and company

★★★★★ Mon · Hinge is sadistic

This company, The Match Group, with their extremely UNETHICAL AND INHUMANE banning policies, are going to create a Luigi Mangione copycat. Let me be clear, it is NOT going to be me, and I am praying that it will not happen, seriously. But this company is literally playing games with the most important aspect of people's lives.

You cannot have a worldwide monopoly in the online dating world and continue to permanently ban/permanently destroy people's lives, and expect to face no consequences. Are you serious? You are LITERALLY DESTROYING PEOPLES LIVES.

At a time when I need support, a time when I need love and companionship because I am struggling from a looming unavoidable death in the family, a time when I need love and companionship THE MOST,

But instead of being able to find it on Hinge or Tinder, they have permanently banned me instead.

This company does NOT care about people's mental health. They WILL permanently ban you and destroy your life.

EMERGING THEMES ACROSS DATING APP REVIEWS



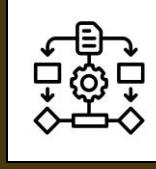
Scams & Fake Profiles

High volume of complaints about catfishing, bots, extortion, and fraudulent accounts.



Safety Concerns

Users express fear around harassment, stalking, and unsafe interactions.



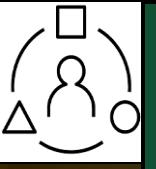
Algorithm Frustrations

Many users feel the matching system is inconsistent, biased, or ineffective.



Monetization Complaints

Strong opinions on expensive subscriptions, paywalls, and "money-hungry" features.



User Experience Failures

Reports on bans with no explanation, glitches, and poor customer support.



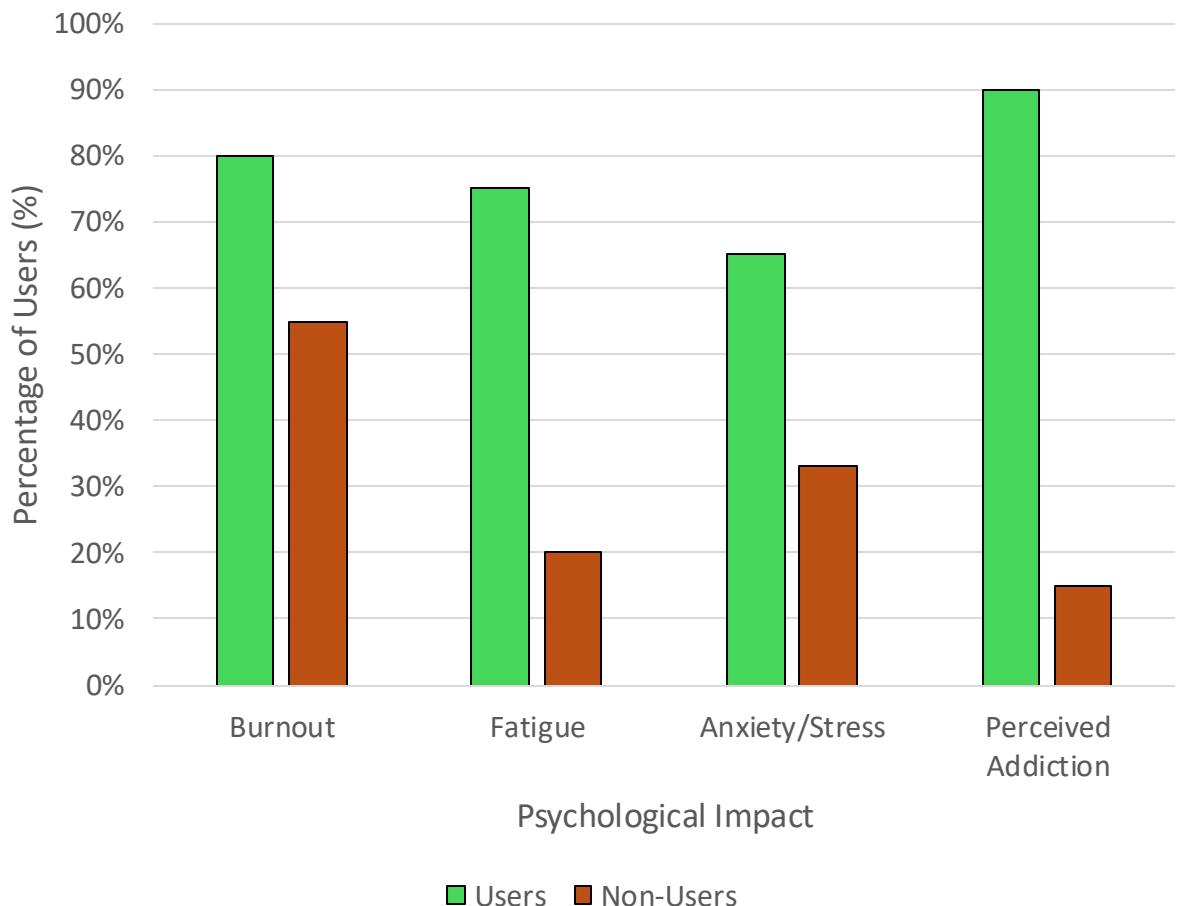
Psychological Fatigue

Burnout, overwhelm, and loss of confidence from repeated negative experiences.

MEASURED PSYCHOLOGICAL IMPACT OF DATING APP USE

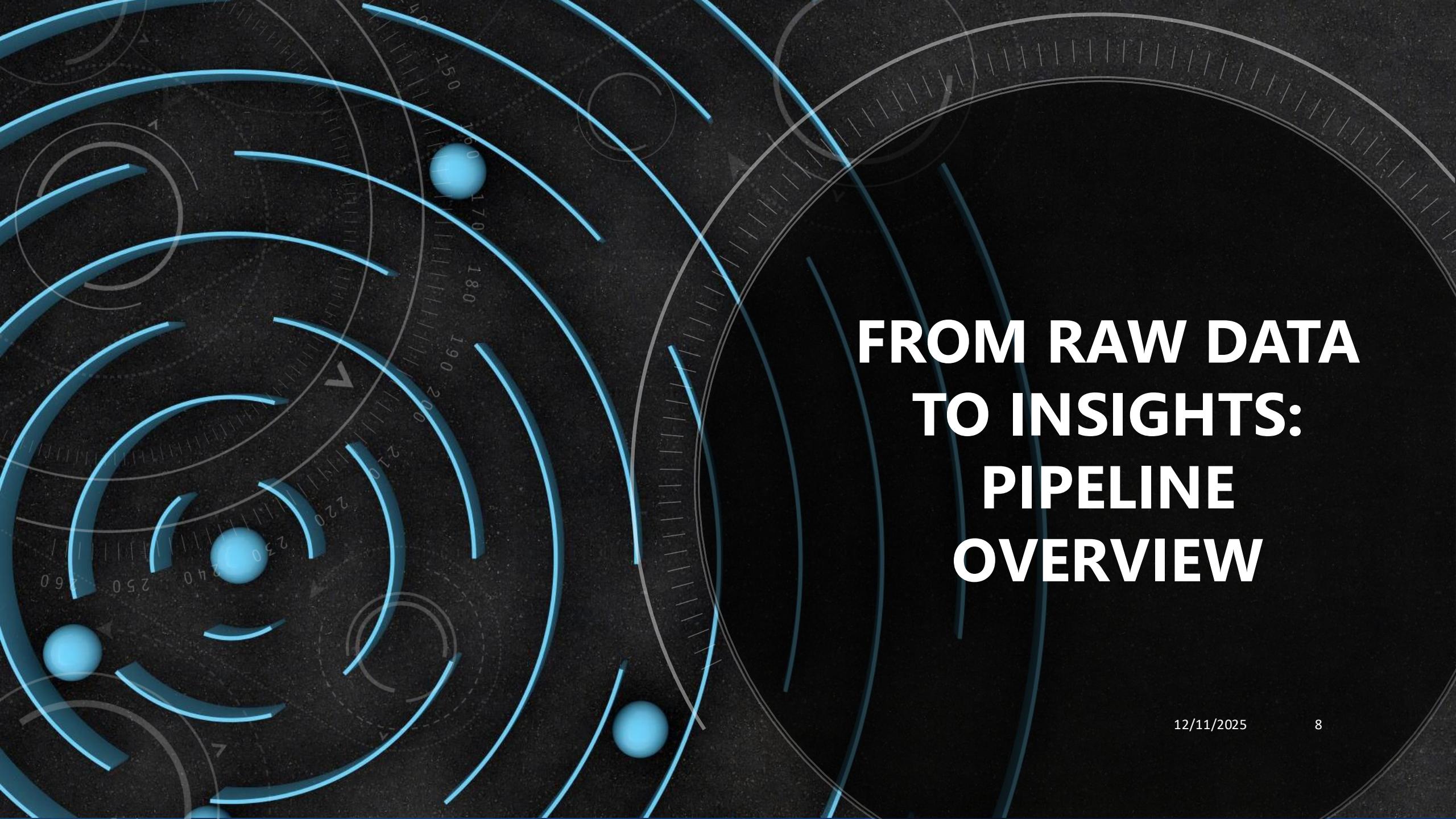
- Users report **2-4x higher levels** of burnout, fatigue, anxiety, and compulsive behavior than non-users.
- Non-users show **significantly lower** emotional distress across all measures.
- The distress gaps reflect **platform-driven stress**, not the mental health baseline.

Psychological Impact of Dating Apps



KEY FINDINGS SO FAR

- User dissatisfaction across dating apps is **systemic**, not just dating behavior.
- Safety, fraud, unknown bans and algorithm instability are the most reported issues.
- Psychological burnout is **structurally induced by platform design**.
- Monetization pressure intensifies negative behavioral outcomes.
- These findings directly shaped the architecture of the governance-first data pipeline, moving beyond surface level observation.



FROM RAW DATA TO INSIGHTS: PIPELINE OVERVIEW

swipe-right-on-data

Summary

ARN: arn:aws:iam::288096952745:role/swipe-right-on-data

Last activity: 22 minutes ago

Permissions

Trust relationships

Tags

Last Accessed

Revoke sessions

Permissions policies (5)

You can attach up to 10 managed policies.

Policy name	Type	Attached entities
AmazonAthenaFullAccess	AWS managed	3
AmazonS3FullAccess	AWS managed	5
AWSGlueConsoleFullAccess	AWS managed	3
AWSGlueServiceRole	AWS managed	4
SwipeRightPolicy	Customer managed	1

Permissions boundary (not set)

Objects (6)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Name	Type	Last modified	Size	Storage class
glue-temp/	Folder	-	-	-
logs/	Folder	-	-	-
parquet/	Folder	-	-	-
processed/	Folder	-	-	-
raw/	Folder	-	-	-
Unsaved/	Folder	-	-	-

AWS IAM ROLE & S3 BUCKET SETUP



Created a custom IAM role with S3, Glue, and Athena permissions to support the entire ETL pipeline.



Set up my S3 bucket structure and added five folders needed for the project.



Configured a glue-temp and logs folder for crawlers, job outputs, and workflow tracking.



Raw folder stores all unprocessed dating app screenshots (Bumble, Hinge, Tinder)



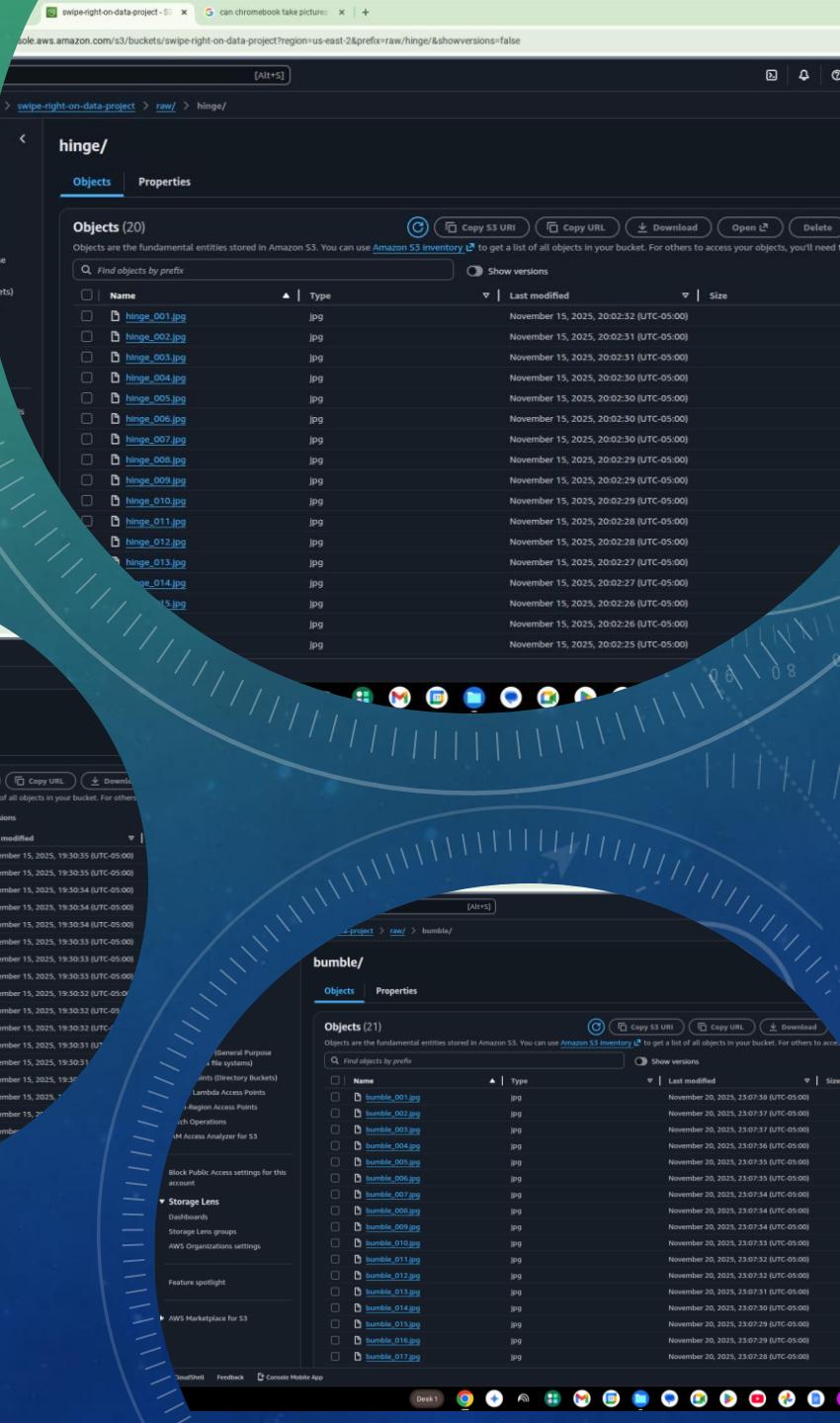
Processed folder stores the cleaned CSV files after Amazon Textract and conversion.



Parquet folder stores optimized files generated for faster Athena queries and cleaner schema management.

DATA INGESTION ARCHITECTURE

- **63 raw screenshots** uploaded to Amazon S3 (Raw Layer)
- Files renamed for consistent formatting.



Step 1: Upload your documents
Upload up to 150 documents, from either your local computer or an S3 bucket. Documents can be multi-page.

Step 2: Choose features
Select one of Textract's features to run on your documents. Note: Bulk Document Uploader usage is charged the same as regular Textract usage.

Step 3: Process documents
Submit a request for Textract to process the documents.

Step 4: Download output
Download a human-readable CSV of the output.

Step 5: Use API in production
Once you are satisfied with the results, use our API for running production workloads.

Output files will remain available for download for 7 days after the completion of processing.

Submitted documents (0/85) Info

Name	Status	Upload date	Type	Feature	Size
image_012	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	411.57 KB
image_006	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	116.14 KB
image_202	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	100.2 KB
image_205	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	81.28 KB
image_016	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	452.72 KB
image_011	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	413.83 KB
image_218	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	340.04 KB
image_001	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	270.0 KB
image_213	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	425.98 KB
image_217	Ready for Download	November 16, 2025, 14:22 (UTC-5:0)	jpg	DetectDocumentText - OCR	356.57 KB

Submitted documents (0/85) Info

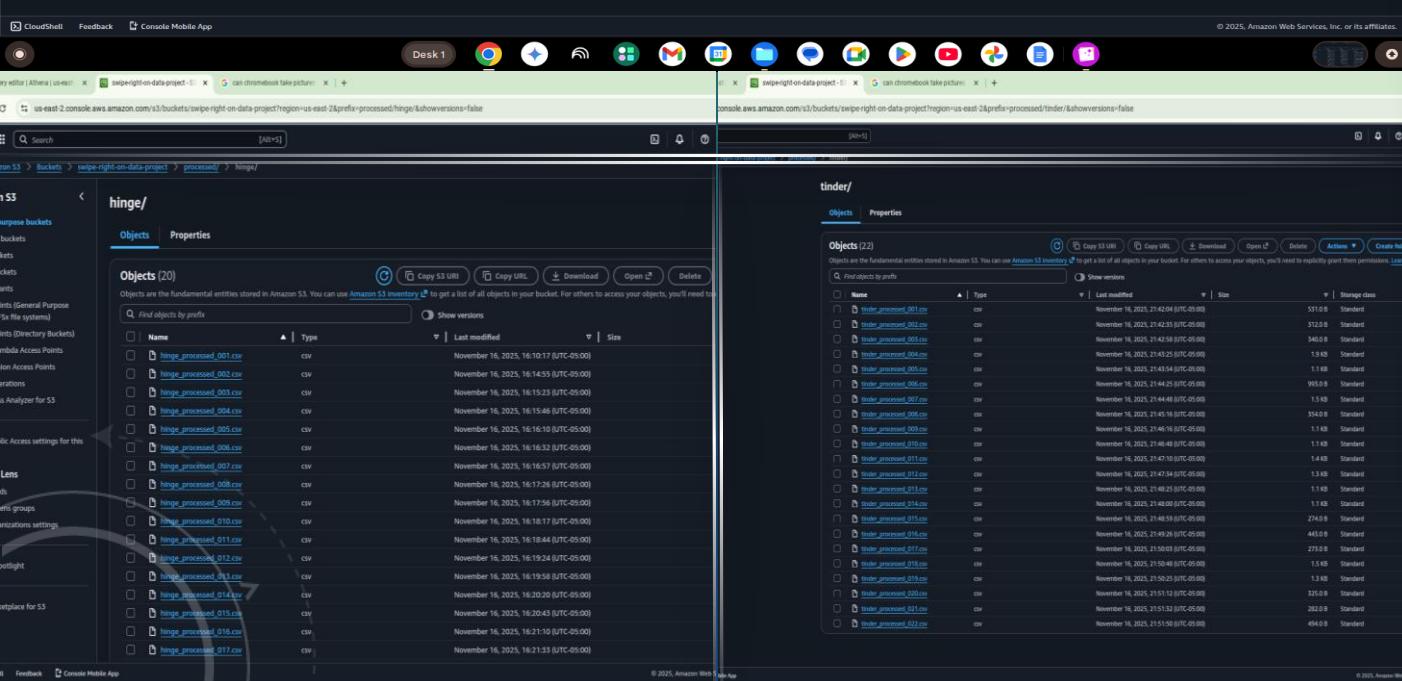
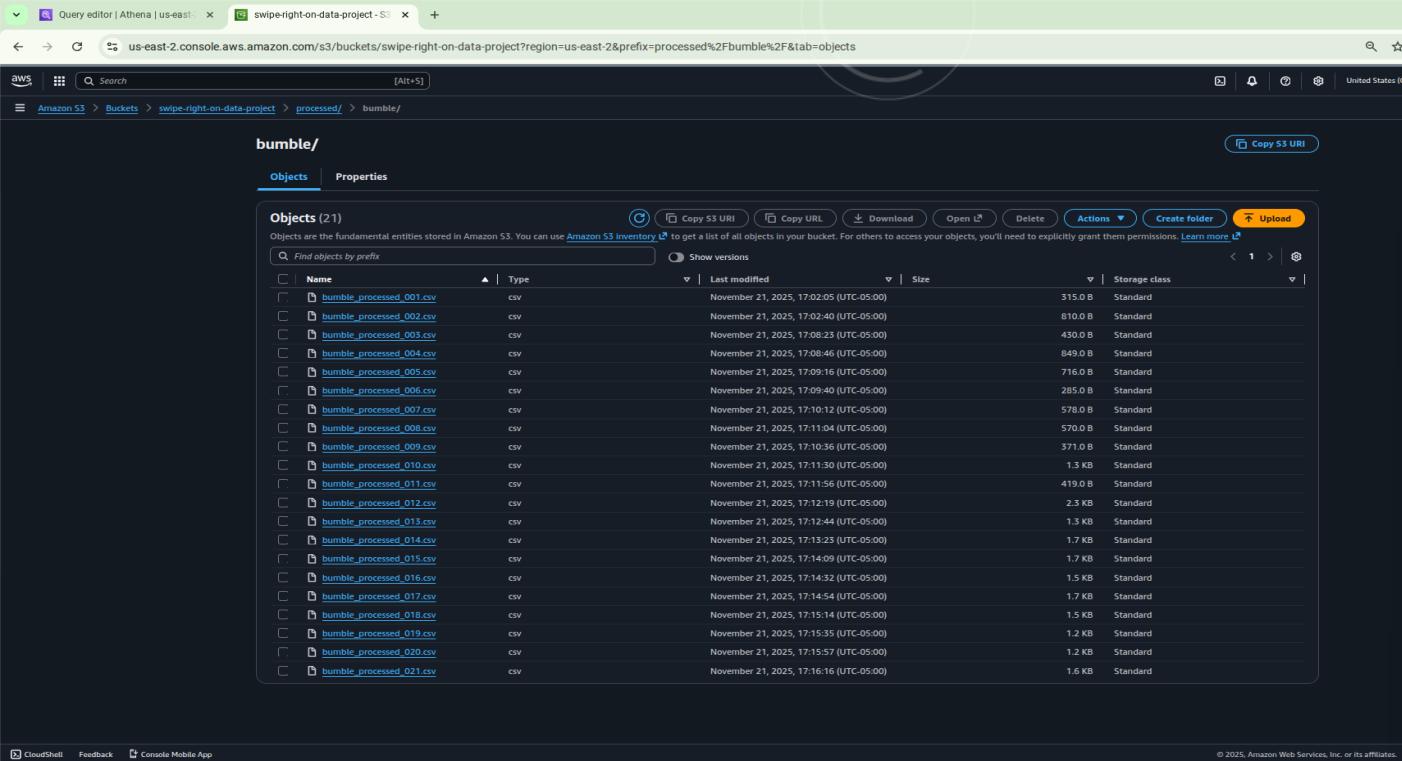
Name	Status	Upload date	Type	Feature	Size
order_014	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	319.59 KB
order_004	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	75.75 KB
order_002	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	169.72 KB
order_001	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	85.13 KB
order_000	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	330.58 KB
order_009	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	104.94 KB
order_006	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	156.68 KB
order_018	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	317.12 KB
order_016	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	100.86 KB
order_012	Ready for Download	November 16, 2025, 16:51 (UTC-5:0)	jpg	DetectDocumentText - OCR	370.14 KB

Submitted documents (0/85) Info

Name	Status	Upload date	Type	Feature	Size
bumble_001	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	376.51 KB
bumble_019	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	334.54 KB
bumble_018	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	414.95 KB
bumble_009	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	404.5 KB
bumble_003	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	279.36 KB
bumble_004	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	377.6 KB
bumble_012	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	621.2 KB
bumble_010	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	132.75 KB
bumble_005	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	199.23 KB
bumble_008	Ready for Download	November 20, 2025, 23:08 (UTC-5:0)	jpg	DetectDocumentText - OCR	184.46 KB

DATA INGESTION ARCHITECTURE

- **Batch processed all raw screenshots using Amazon Textract.**
- **Extracted and converted unstructured images into structured CSV files.**
- **Reviewed Amazon Textract output to ensure clean extraction to text from each dating-app review.**
- **Processed CSV files stored in S3 (Processed Layer).**



DATA INGESTION ARCHITECTURE

- **Uploaded all structured CSV files to the S3 processed folder.**
 - **Files renamed** for consistent formatting and automated schema detection in AWS Glue.

Announcing new optimization features for Apache Iceberg tables
Optimize storage for Apache Iceberg tables with automatic snapshot retention and orphan file deletion. [Learn more](#)

swipe_right_on_data

Database properties

Name: swipe_right_on_data

Description:

Location: s3://swipe-right-on-data-project

Tables (8)

View and manage all available tables.

Filter tables

Name	Database	Location	Classification	Deprecated
bumble_processed	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-
bumble_processedbumble	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-
dating_app_combined_data	swipe_right_on_data	s3://swipe-right-on-data-project/prc	Parquet	-
hinge_processed	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-
hinge_processedhinge	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-
parquet	swipe_right_on_data	s3://swipe-right-on-data-project/prc	Parquet	-
tinder_processed	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-
tinder_processedtinder	swipe_right_on_data	s3://swipe-right-on-data-project/prc	CSV	-

Crawlers (4)

A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.

Name	State	Schedule	Last run	Last run timestamp	Log	Table count
bumble-crawler	Ready		Succeeded	November 21, 2025 at 06:24:02	View log	1 created
hinge-crawler	Ready		Succeeded	November 16, 2025 at 21:31:42	View log	1 created
parquet-crawler	Ready		Succeeded	November 21, 2025 at 21:40:36	View log	1 created
tinder-crawler	Ready		Succeeded	November 17, 2025 at 02:55:58	View log	1 created

AWS GLUE & SCHEMA MANAGEMENT

- Databases were created to organize all structured review data.
- AWS Glue Crawlers detected schemas from Textract CSV outputs.
- Verified table structures in Data Catalog ensured that it was a clean query performance.

Table details

Schema (4)

#	Column name	Data type	Partition key	Comment
1	page number	bigint	-	-
2	type	string	-	-
3	text	string	-	-
4	confidence score % (line)	double	-	-

Table details

Schema (4)

#	Column name	Data type	Partition key	Comment
1	page number	bigint	-	-
2	type	string	-	-
3	text	string	-	-
4	confidence score % (line)	double	-	-

Table details

Schema (4)

#	Column name	Data type	Partition key	Comment
1	page number	bigint	-	-
2	type	string	-	-
3	text	string	-	-
4	confidence score % (line)	double	-	-

Raw Schema: Pre-Transformation View

- Schemas were **validated for column accuracy, page numbers, text, types and confidence scores** to avoid duplication and manipulation.
- All tables were registered into the Glue Data Catalog under "swipe_right_on_data"

The screenshot shows the AWS Athena Query Editor interface. On the left, the sidebar displays the Data source (AwsDataCatalog), Catalog (None), Database (swipe_right_on_data), and Tables (0). The main area contains a SQL query: `SELECT * FROM "swipe_right_on_data"."hinge_processedhinge" limit 300;`. Below the query, the status bar indicates "SQL Ln 1, Col 68". At the bottom of the editor, there are buttons for "Run again", "Explain", "Cancel", "Clear", and "Create". The "Query results" tab is selected, showing the results of the completed query. The results table has columns: #, page number, type, text, and confidence score % (line). The results show various text snippets from the hinge_processedhinge table, such as "I met HIM", "Mon Ajee M", and "I never thought I would meet someone who was simply created with me in mind until I met him.". The run time is listed as 411 ms.

This screenshot shows two separate instances of the AWS Athena Query Editor running side-by-side. Both instances have the same basic structure: a sidebar with a Data source (AwsDataCatalog), Catalog (None), Database (tinder_processed), and Tables (0). The top instance has a query in progress: `SELECT * FROM "swipe_right_on_data"."tinder_processedtinder" limit 300;`. The bottom instance also has a query in progress: `SELECT * FROM "swipe_right_on_data"."hinge_processedhinge" limit 300;`. Both instances show the "Query results" tab with tables of results. The results for the top query include rows like "Jose Olvera", "X.com", and "@OliverOlvera". The results for the bottom query include rows like "Abe", "X.com", and "@gabe_xxx". The run times for both queries are listed as 360 ms.

ATHENA QUERY VALIDATION

- SQL queries **validated raw and processed schemas.**
- Sample records reviewed for consistency and data accuracy.
- Row counts **verified** per platform.
- Data is ready for **unified** analysis!

Google Chrome tab: can chromebook take pictures

Visual - Editor - AWS Glue Studio tab: d982e275012e12d23d5/MyFiles/Downloads/Visual%20-%20Editor%20-%20AWS%20Glue%20Studio.mhtml

Amazon CloudWatch Metrics tab: console.aws.amazon.com/gluestudio/home?region=us-east-2#/editor/job/Swipe%20Right%20On%20Data/runs

The screenshot shows the AWS Glue Studio visual editor on the left and the CloudWatch Metrics console on the right.

AWS Glue Studio (Left):

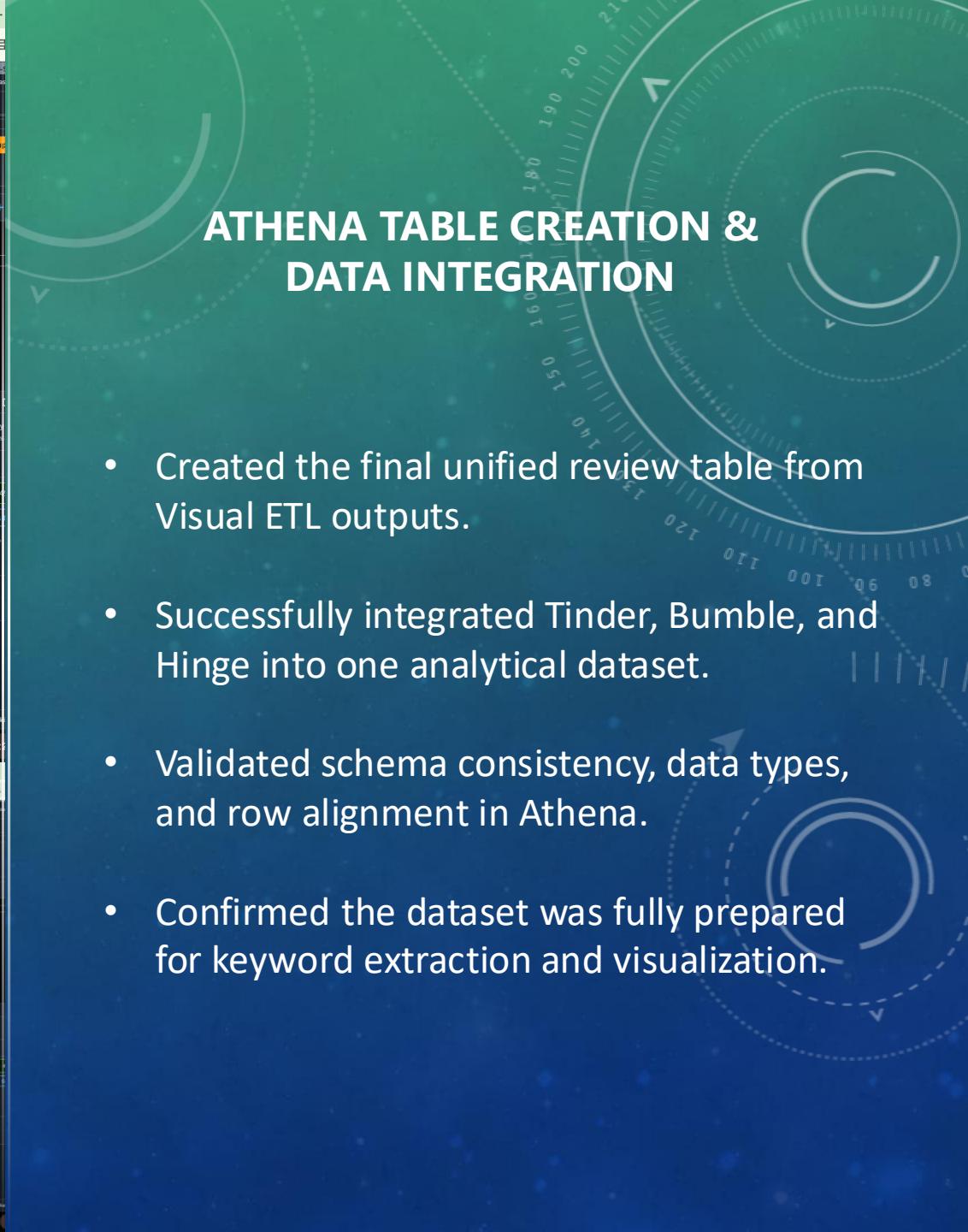
- Screenshots:** Visual, Script, Job details, **Runs**, Data quality, Schedules, Version Control.
- Job Flow Diagram:** A directed acyclic graph (DAG) showing the data processing pipeline.
 - Data Sources:** Three S3 buckets: "Bumble S3", "Hinge S3", and "Tinder S3".
 - Transformations:** Two "Change Schema" transforms (one for Bumble and Hinge S3) and two "Join Outer Join" transforms (one for each pair of source datasets).
 - Data Target:** One S3 bucket named "joined_dating_output".
 - Connections:** Bumble S3 → Change Schema (Bumble) → Join Outer Join (Bumble, Hinge) → joined_dating_output. Hinge S3 → Change Schema (Hinge) → Join Outer Join (Bumble, Hinge) → joined_dating_output. Tinder S3 → Change Schema (Tinder) → Join Outer Join (Tinder, Hinge) → joined_dating_output.

CloudWatch Metrics (Right):

- Section:** Swipe Right On Data
- Job runs (1/4):** Info
- Table:** Shows four job runs with the following details:

Run status	Retries	Start time (Local)	End time (Local)	Duration
Succeeded	0	11/21/2025 03:03:19	11/21/2025 03:04:28	1 m 3 s
Succeeded	0	11/21/2025 02:28:22	11/21/2025 02:29:38	1 m 9 s
Failed	0	11/21/2025 02:24:46	11/21/2025 02:25:45	52 s
Failed	0	11/21/2025 02:19:05	11/21/2025 02:20:05	52 s
- Run details:** Tab selected. Shows the following information:

Job name	Start time (Local)	Glue version
Swipe Right On Data	11/21/2025 03:03:19	5.0
Id	End time (Local)	Worker type
jr_a7033562e740b8e8e95b4e8c3a5c7d6d42f72df5bf2ae9920b3bf25b6e2346	11/21/2025 03:04:28	G.1X
Run status	Start-up time	Max capacity
Succeeded	6 seconds	10 DPU
Retry attempt number	Execution time	Execution class
Initial run	1 minute 3 seconds	Standard
Trigger name	Security configuration	Cloudwatch logs
- Input arguments (9):** Shows nine input arguments with their values.
- Logs:** Shows log entries.
- Run insights:** Shows run insights.
- Metrics:** Shows metrics.
- Troubleshooting analysis - preview:** Shows troubleshooting analysis.
- Spark UI:** Shows Spark UI details.



ATHENA TABLE CREATION & DATA INTEGRATION

- Created the final unified review table from Visual ETL outputs.
- Successfully integrated Tinder, Bumble, and Hinge into one analytical dataset.
- Validated schema consistency, data types, and row alignment in Athena.
- Confirmed the dataset was fully prepared for keyword extraction and visualization.

Query in Amazon SageMaker Unified Studio

Use your IAM role to analyze and build with your existing AWS resources in a single data and AI development environment.

Query 68 [Alt+5]

```
1 SELECT source_app, COUNT(*) AS review_count
2 FROM "swipe_right_on_data"."parquet"
3 GROUP BY source_app
4 ORDER BY review_count DESC;
5
```

SQL Ln 4, Col 28

Run again Explain Cancel Clear Create

Query results | query stats

Completed

Time in queue: 72 ms Run time: 449 ms Data scanned: 3.4 GB

Results (2)

#	source_app	review_count
1	LINE	164835
2		341

Copy Download results

Query in Amazon SageMaker Unified Studio

Analyze and build with your existing AWS resources in a single data and AI development environment.

Query 69 [Alt+5]

```
1 SELECT
2     source_app,
3     AVG(Confidence) AS avg_confidence,
4     COUNT(*) AS review_count
5     FROM "swipe_right_on_data"."parquet"
6     GROUP BY source_app
7     ORDER BY avg_confidence DESC;
```

SQL Ln 3, Col 39

Run again Explain Cancel Clear Create

Query results | query stats

Completed

Time in queue: 88 ms Run time: 31 ms Data scanned: 3.4 GB

Results (2)

#	source_app	avg_confidence
1	LINE	97.9
2		341

Query editor | Athena | us-east-2 | Query editor | Athena | us-east-2 | us-east-2.console.aws.amazon.com/athena/home?region=us-east-2#/query-editor/history/40e9f308-afe8-45ed-9a1c-f9e70843ff01

Account ID: 2880-9693-2745 | DashPenington | United States (Ohio) | +

Amazon Athena > Query editor

Data source: AwsDataCatalog

Catalog: None

Database: swipe_right_on_data

Tables and views:

- Tables (8)
 - bumble_processed
 - bumble_processedbumble
 - dating_app_combined_data
 - hinge_processed
 - hinge_processedhinge
 - page number
 - type
 - text
 - confidence score % (line)
 - parquet
 - page_number_tinder
 - source_app
 - review_text
 - confidence
 - page_number_bumble
 - page_number_hinge
 - tinder_processed
 - tinder_processedtinder
- Views (0)

SQL: SELECT * FROM "swipe_right_on_data"."parquet" limit 10;

Ln 1, Col 1

Run again Explain Cancel Clear Create Reuse query results up to 60 minutes ago

Query results: Completed Time in queue: 94 ms Run time: 350 ms Data scanned: 3.82 KB

Results (10)

#	page_number_tinder	source_app	review_text	confidence	page_number_bumble	page_number_hinge
1		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
2		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
3		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
4		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
5		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
6		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
7		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
8		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
9		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1
10		LINE	6:59 AM 11/7/25 1.4K Views	99.47855377	1	1

CloudShell Feedback Console Mobile App © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Desk 1 Nov 21 5:20

us-east-2.console.aws.amazon.com/athena/home?region=us-east-2#/query-editor/history/bbd2792a-f1fb-40ec-996d-b3c0eeef4c41

Query editor | Athena | us-east-2

[Alt+S]

Query 63 :

```

1 select
2 t.text as tinder_text,
3 b.text as bumble_text,
4 h.text as hinge_text,
5
6 t."type" as tinder_type,
7 b."type" as bumble_type,
8 h."type" as hinge_type
9
10 from tinder_processedtinder t
11 full outer join bumble_processedbumble b
12 on t.text = b.text
13 full outer join hinge_processedhinge h
14 on coalesce(b.text, h.text) = t.text
15 LIMIT 50;

```

SQL Ln 15, Col 10

Run again Explain Cancel Clear Create

Query results | **Query stats**

Completed

Results (50)

#	tinder_text	bumble_text	hinge_text
1	PitonY		
2	X.com	X.com	have to subscribe. But the free version works jus
3	X.com	X.com	premium features and seeing your matches you
4	X.com	X.com	"what really separates this app from others. Yes
5	X.com	X.com	customization and prompts/dating intentions is
6	X.com	X.com	"dating apps
7	X.com	X.com	I will start this off by saying that out of all the
8	X.com	X.com	Mon Connor2208
9	X.com	X.com	The best dating app I found my one!
10	X.com	X.com	Tinder where you just swipe right or left on a per

CloudShell Feedback Console Mobile App

aws Search [Alt+S]

Amazon Athena > Query editor

Data

Data source: AwsDataCatalog

Catalog: None

Database: swipe_right_on_data

Tables and views

Tables (8): bumble_processed, bumble_processedbumble, dating_app_combined_data, hinge_processed, hinge_processedhinge, parquet, tinder_processed, tinder_processedtinder

Views (0)

Query 63 :

```

1 select
2 t.text as tinder_text,
3 b.text as bumble_text,
4 h.text as hinge_text,
5
6 t."type" as tinder_type,
7 b."type" as bumble_type,
8 h."type" as hinge_type
9
10 from tinder_processedtinder t
11 full outer join bumble_processedbumble b
12 on t.text = b.text
13 full outer join hinge_processedhinge h
14 on coalesce(b.text, t.text) = h.text
15 LIMIT 50;

```

SQL Ln 14, Col 32

Run again Explain Cancel Clear Create

Query results | **Query stats**

Completed

Results (50)

#	tinder_text
1	Fixed in no time!
2	Aug 15 . GoldenHoneybee777
3	So around 3-4pm today my account got locked
4	out(i have no clue why) and it wanted me to verify
5	that it was me. Well in basically NO TIME tinder
6	fixed it and im able to log back in! Its now 6pm so
7	i am incredibly thankful on how fast they were able
8	to resolve the issue!!
9	John
10	Jul 13 . anniemagic3

CloudShell Feedback Console Mobile App

Desk 1

Amazon Athena > Query editor

Query 63 :

```
1 select
2 coalesce(t.text, b.text, h.text) as unified_text,
3 t.text as tinder_text,
4 b.text as bumble_text,
5 h.text as hinge_text,
6
7 t."type" as tinder_type,
8 b."type" as bumble_type,
9 h."type" as hinge_type
10
11 from tinder_processedtinder t
12 full outer join bumble_processedbumble b
13 on t.text = b.text
14 full outer join hinge_processedhinge h
15 on coalesce(t.text, b.text) = h.text;
```

SQL Ln 15, Col 38

Run again Explain Cancel Clear Create

Query results | Query stats

Completed

Results (1,946)

#	unified_text	tinder_text	bumble_text	hinge_text	tinder_type	bumble_type	hinge_type
1901	every time i open the app im like			every time i open the app im like			LINE
1902	tz 12			tz 12			LINE
1903	Mon . Mac39210			Mon . Mac39210			LINE
1904	Mon . Jenribot			Mon . Jenribot			LINE
1905	and have been through it in relationships and are			and have been through it in relationships and are			LINE
1906	Mon Armando&Aletri			Mon Armando&Aletri			LINE
1907	Hinge has blessed us with the gift of love.			Hinge has blessed us with the gift of love.			LINE
1908	Banned for no reason			Banned for no reason			LINE
1909	am honestly shocked and frustrated with Hinge			am honestly shocked and frustrated with Hinge			LINE

Time in queue: 89 ms Run time: 442 ms Data scanned: 58.26 KB

Copy Download results CSV

CloudShell Feedback Console Mobile App

Desk 1

Amazon Athena > Query editor

Query 63 :

```
1 select
2 t.text as tinder_text,
3 b.text as bumble_text,
4 h.text as hinge_text,
5
6 t."type" as tinder_type,
7 b."type" as bumble_type,
8 h."type" as hinge_type
9
10 from tinder_processedtinder t
11 full outer join bumble_processedbumble b
12 on t.text = b.text
13 full outer join hinge_processedhinge h
14 on coalesce(t.text, b.text) = h.text;
```

SQL Ln 14, Col 38

Run again Explain Cancel Clear Create

Query results | Query stats

Completed

Results (1,946)

#	tinder_text	bumble_text	hinge_text	tinder_type	bumble_type	hinge_type
1301			Introducing the new			
1302			10:06 AM . 5/13/24 2.9M Views			
1303			Cash grab			
1304			Nov 4 . pinkisme7			
1305			barely touched the screen and it glitched and the			
1306			Not for men to choose			
1307			Nov 4 . CooperStayton			
1308			profile was phased out of popularity to never be			
1309			wouldn't get any matches			
1310			supposed to be rejected.			

Time in queue: 94 ms Run time: 597 ms Data scanned: 58.26 KB

Copy Download results CSV

CloudShell Feedback Console Mobile App

Desk 1

ATHENA: VALIDATION & INTEGRATION

- Validated all Glue-generated schemas using Athena SQL queries.
- Integrated Tinder, Bumble, and Hinge via multi-table JOIN operations.
- Verified data volume using COUNT queries (164K+ reviews).
- Computed star ratings and review metrics for visualization.
- Confirmed stable query performance with zero schema or permission errors.

Quick summary

Athena_Visualizationin

37178

Sum of page_nu...

15,626,158.43

Sum of confidence

160000

Count of Athena...

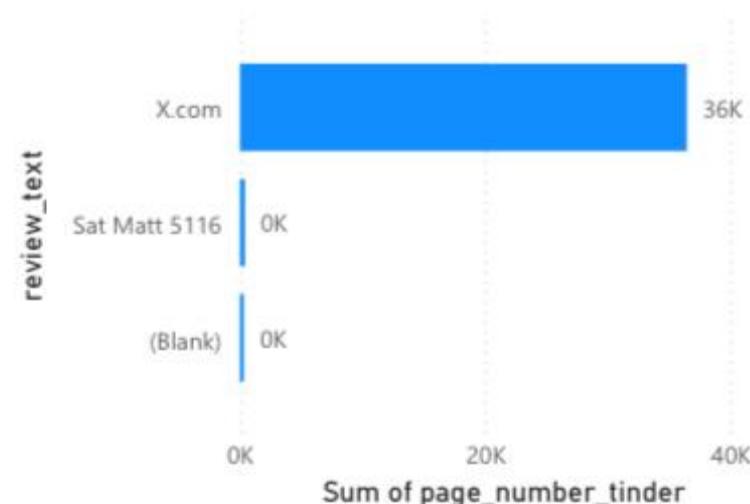
159677

Sum of page_nu...

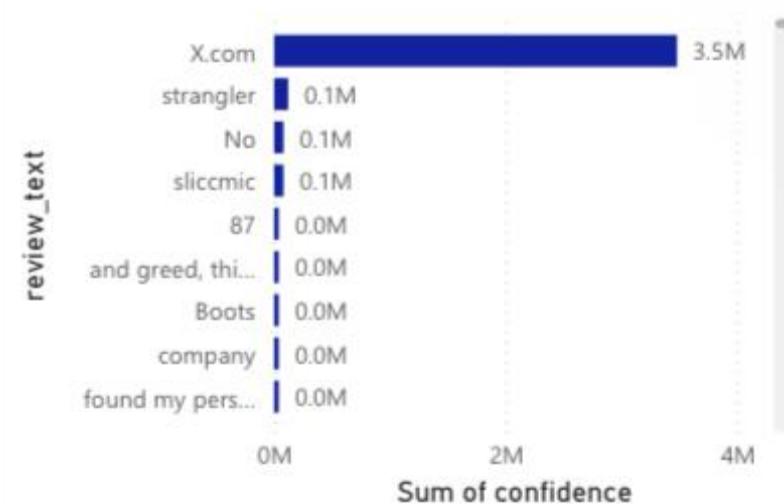
159677

Sum of page_nu...

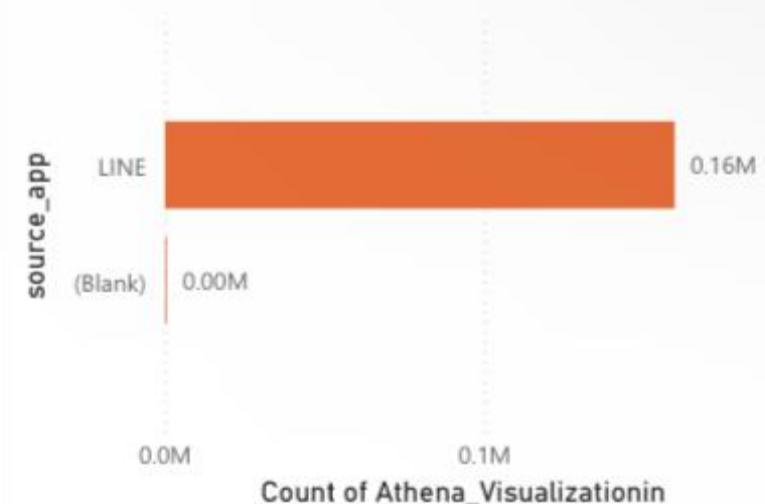
Sum of page_number_tinder by review_text



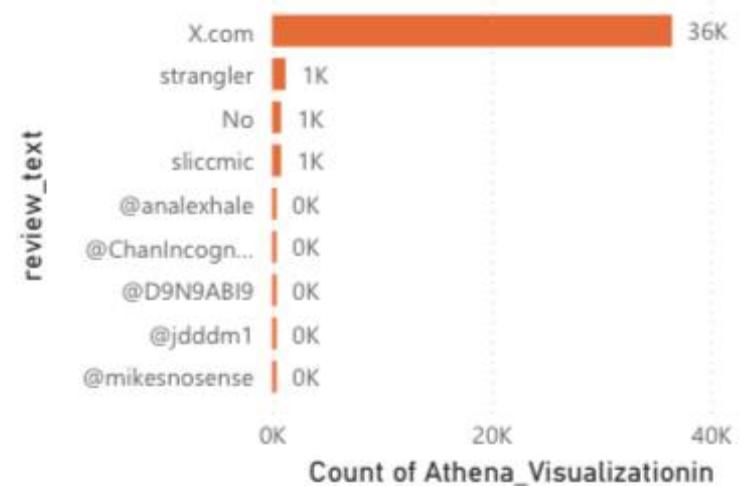
Sum of confidence by review_text



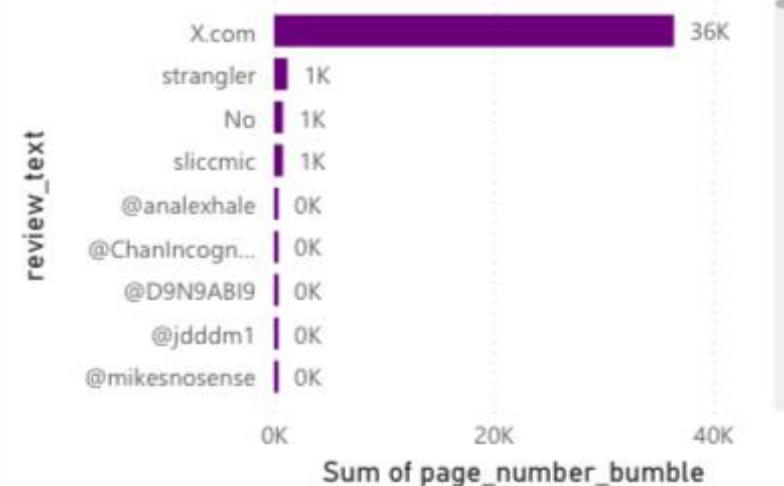
Count of Athena_Visualizationin by source_app



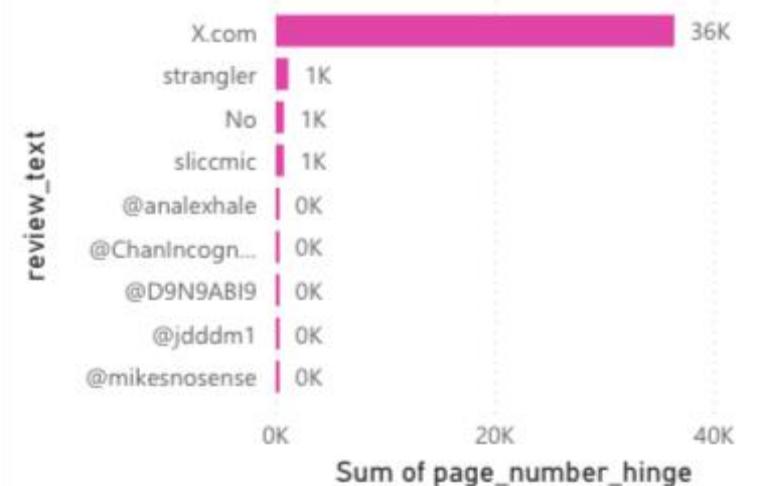
Count of Athena_Visualizationin by review_text



Sum of page_number_bumble by review_text



Sum of page_number_hinge by review_text



Quick summary

Athena_Full

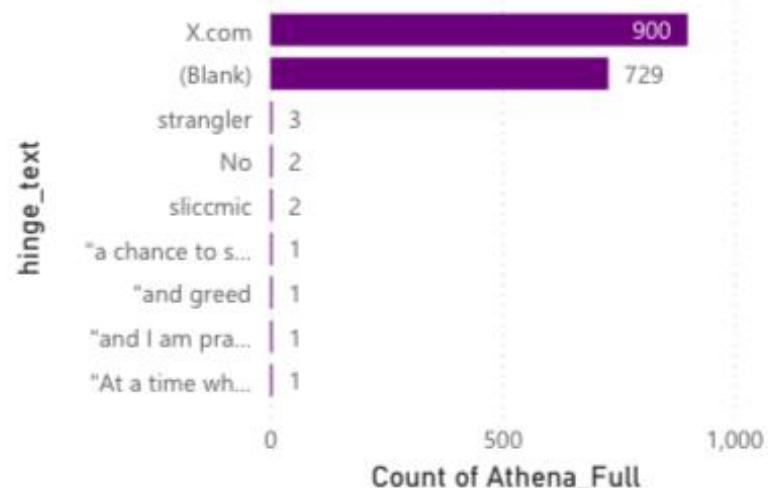
1946

Count of Athena_Full

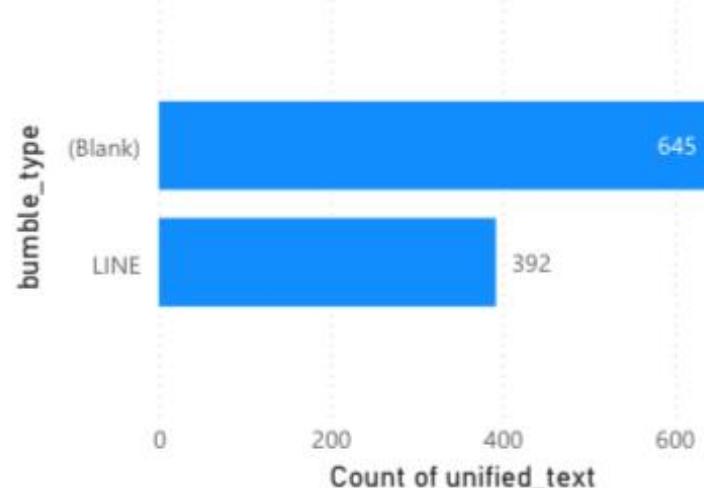
1035

Count of unified_text

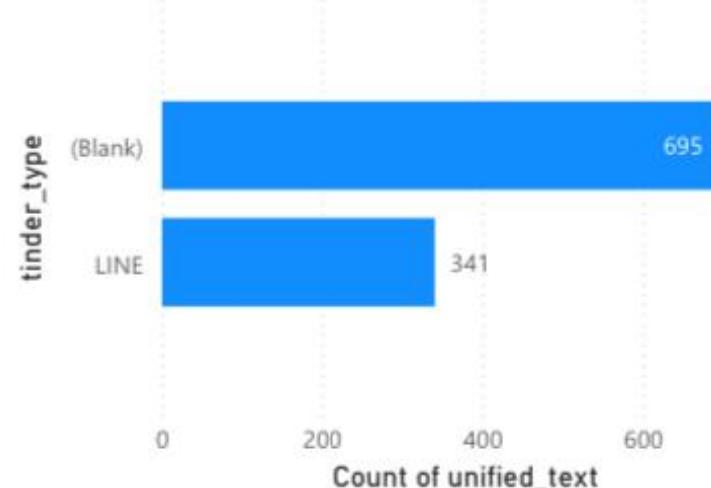
Count of Athena_Full by hinge_text



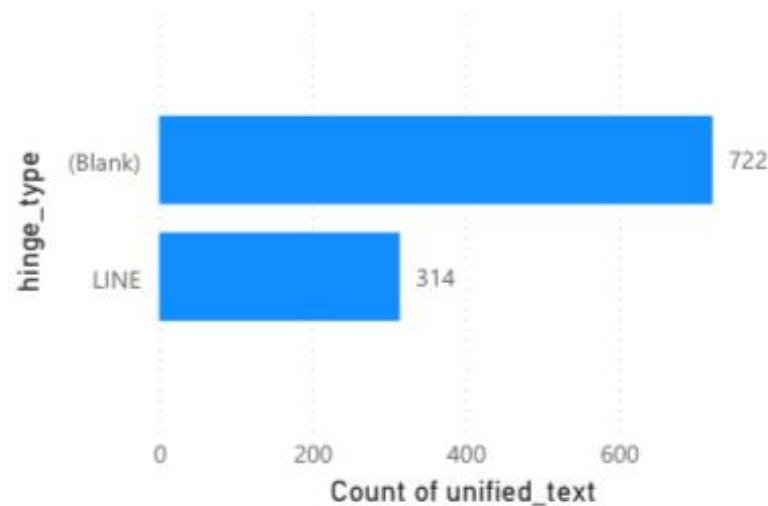
Count of unified_text by bumble_type



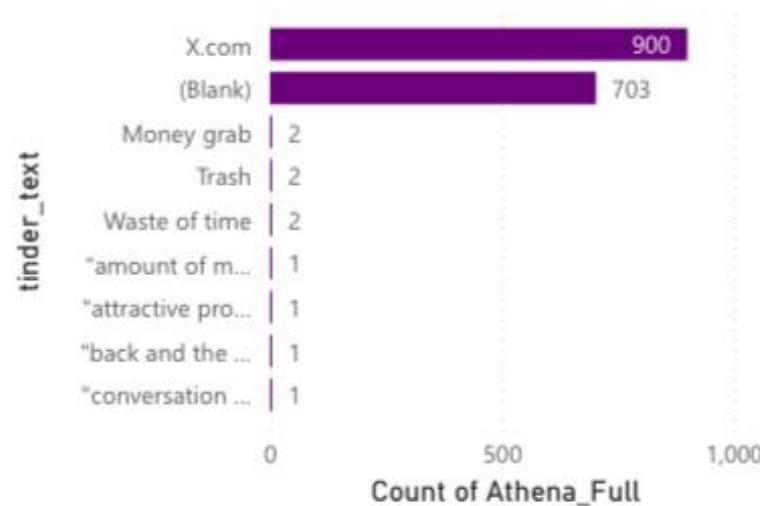
Count of unified_text by tinder_type



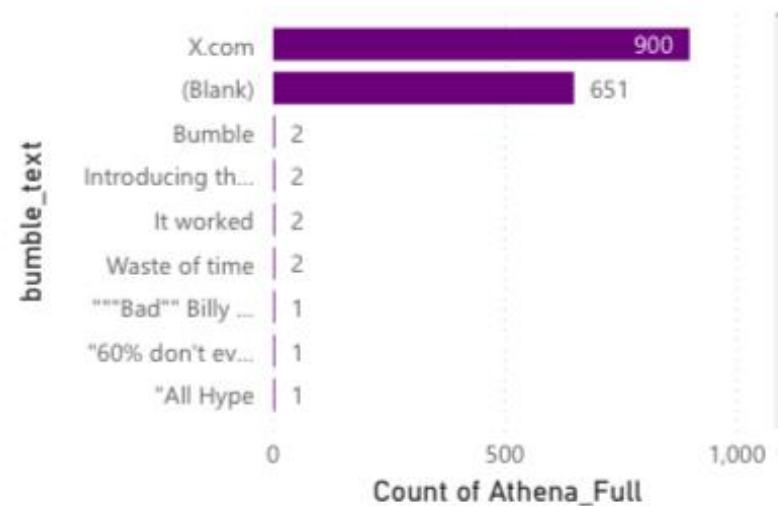
Count of unified_text by hinge_type



Count of Athena_Full by tinder_text



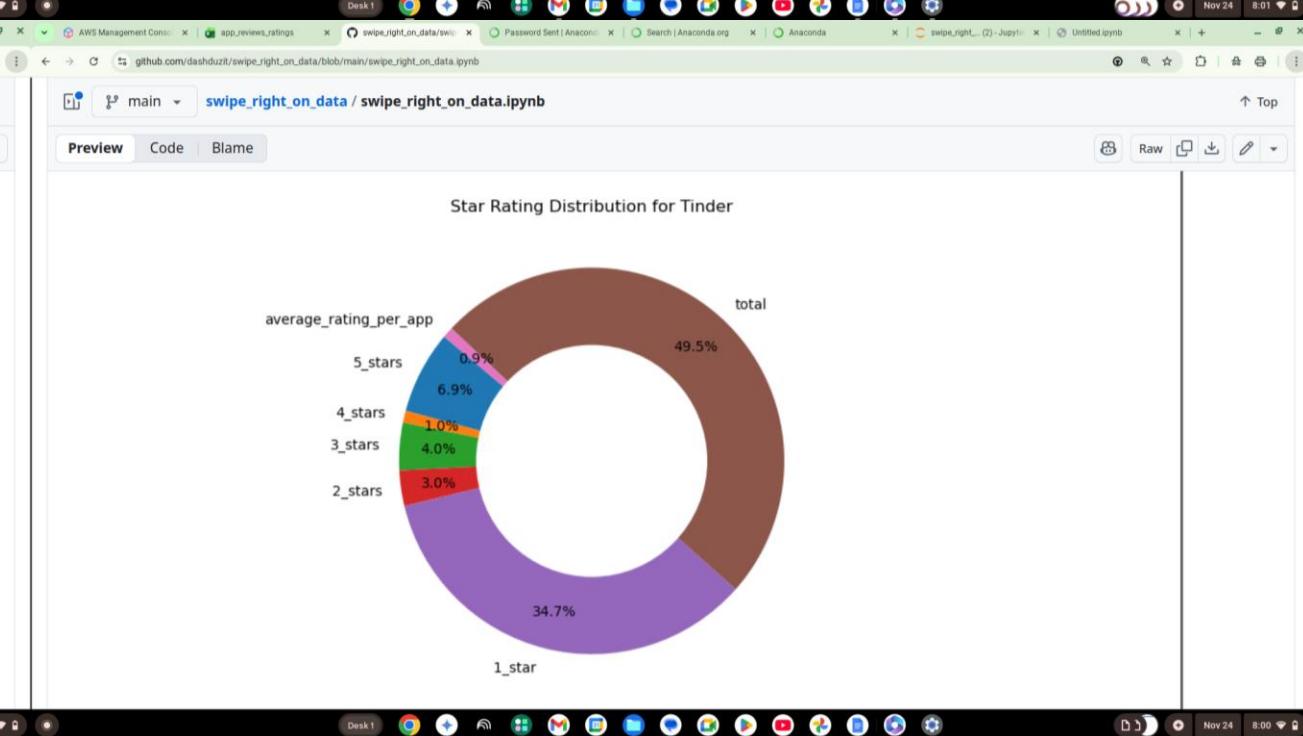
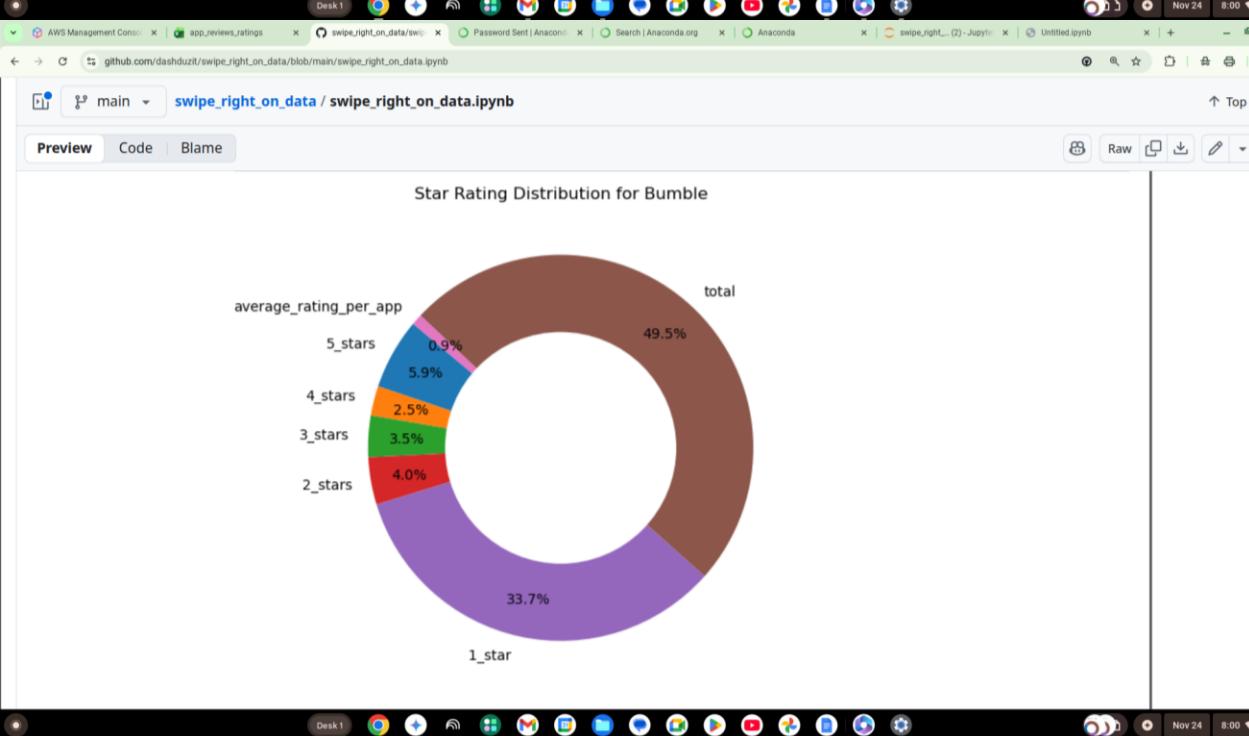
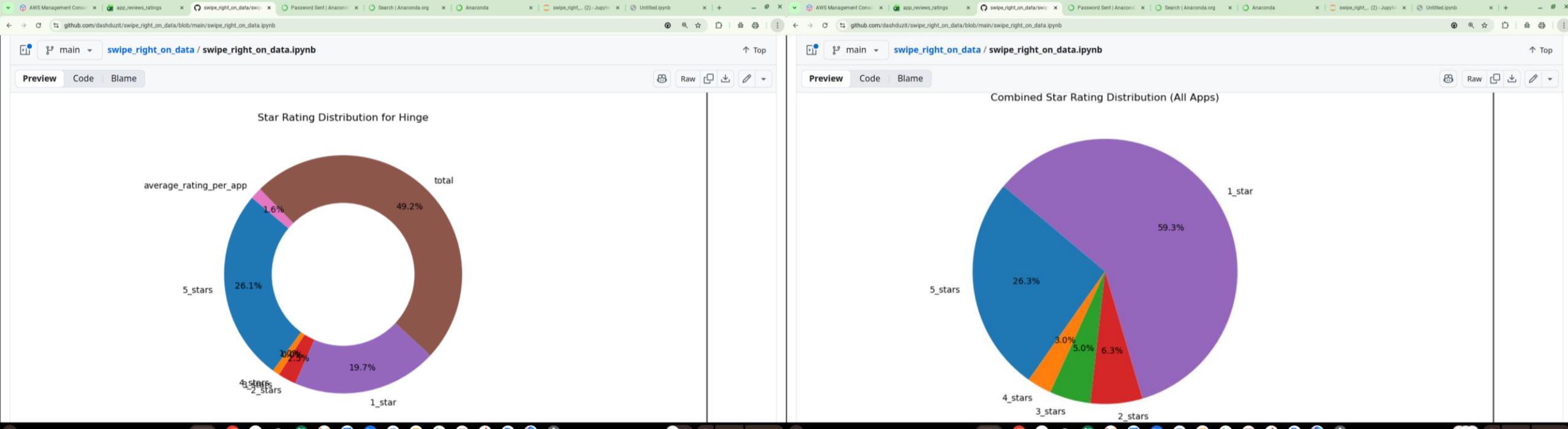
Count of Athena_Full by bumble_text



A man in a dark suit and tie looks directly at the camera with a weary expression. He is positioned in front of a grid of numerous people, each holding up a large amount of US dollar bills. The scene is overlaid with several circular data visualization elements, including radar-like charts and a clock face, suggesting a theme of time and measurement. The overall mood is one of frustration or despair.

**RANKING THE DATING APPS BY
AVERAGE DISAPPOINTMENT**

(BASED ON ACTUAL USER SUFFERING)

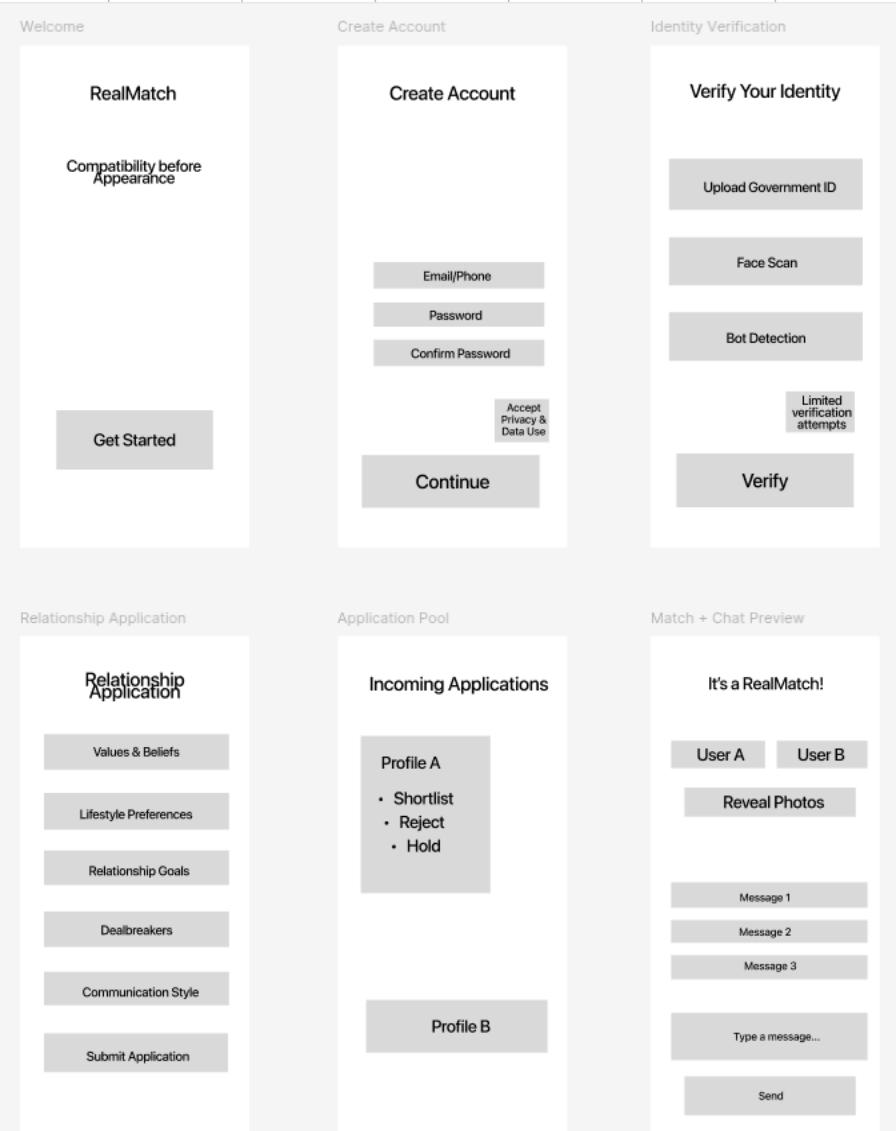
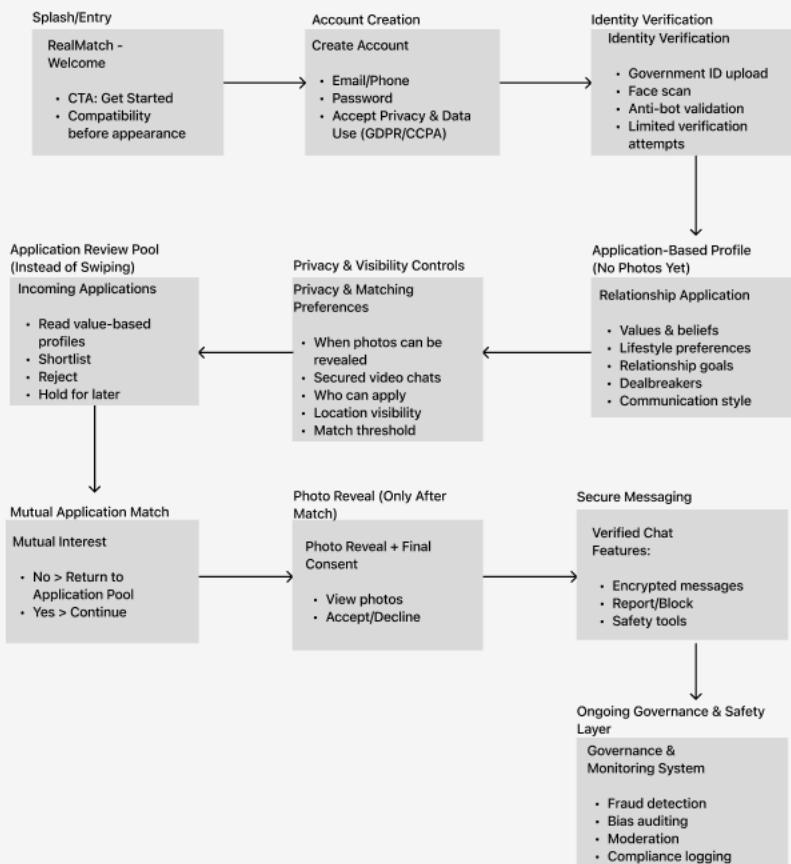




PROPOSED SOLUTION

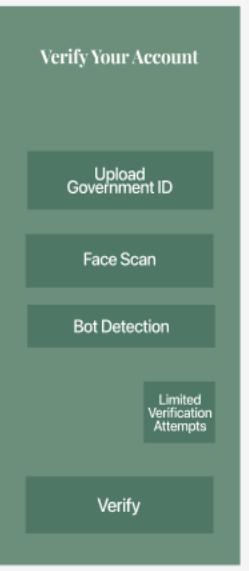
REALMATCH: GOVERNANCE-FIRST DATING PLATFORM

- **User-Centered Design:** Built from real user frustrations across Tinder, Bumble, and Hinge.
- **Job-Application Matching Model:** Encourages intentional selection and reduces impulsive swiping.
- **Identity Verification:** Double authentication to minimize bots and catfishing.
- **Safety & Governance Focus:** Addresses misrepresentations, harassment, and platform trust gaps.
- **Secure Messaging, Private Sharing & Encrypted Media:** Protects user privacy and prevents content misuse with safer and controlled communication aligned with data protection standards.





Relationship Application



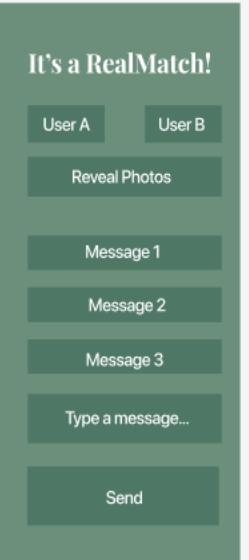
Match + Chat Preview



ANSWER



© 2010 Pearson Education, Inc.



Digitized by srujanika@gmail.com

DATING APPS VS. REALMATCH: KEY COMPARISONS



MATCHING PROCESS: TRADITIONAL APPS RELY ON SWIPE-BASED MATCHING → REALMATCH USES A JOB-APPLICATION STYLE MATCHING MODEL TO PROMOTE FAIRNESS AND CONSISTENT ALGORITHMS.



SAFETY & AUTHENTICATION: OTHER APPS HAVE LIMITED VERIFICATION WHICH CAN BE ABUSED BY BOTS, SCAMMERS, AND CATFISHING → REALMATCH USES DOUBLE AUTHENTICATION TO PROMOTE AUTHENTICITY.



GOVERNANCE & COMPLIANCE: EXISTING PLATFORMS SHOW INCONSISTENT ENFORCEMENT → REALMATCH IS BUILT GDPR & CCPA FIRST TO PROMOTE SAFETY AND TRUST.



MONETIZATION STRUCTURE: PAYWALLS, BOOSTS, MICROTRANSACTIONS WHICH CAN BE HARMFUL AND PREDATORY → REALMATCH USES A MONTHLY SUBSCRIPTION FEE (\$19.99).

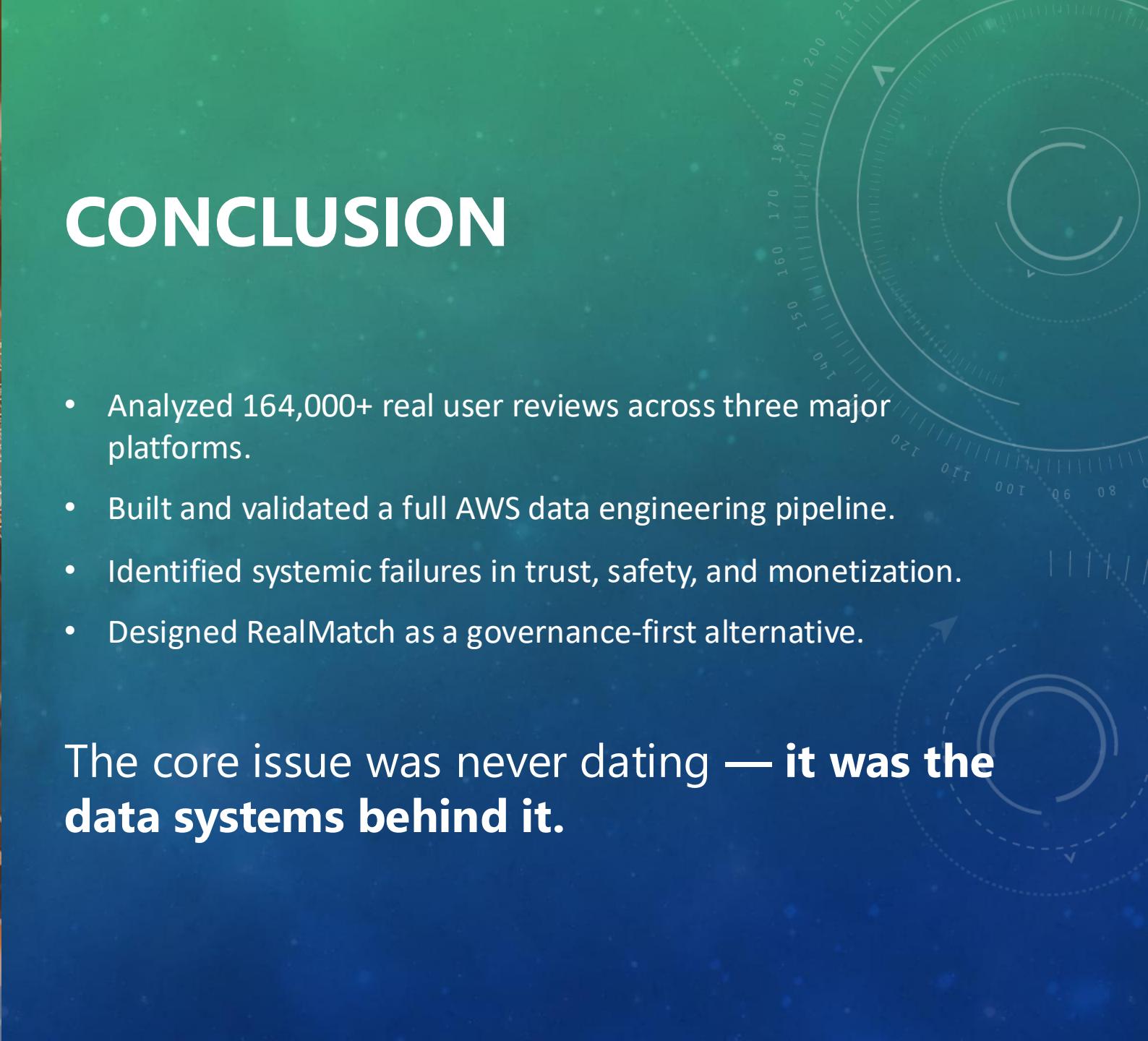


USER EXPERIENCE IMPACT: CURRENT APPS INCREASE BURNOUT AND SUPERFICIAL FILTERING → REALMATCH IS SELF-PACED AND TRUST-CENTERED TO BUILD CONFIDENCE AND HEALTHIER ENGAGEMENT.



CONCLUSION

- Analyzed 164,000+ real user reviews across three major platforms.
- Built and validated a full AWS data engineering pipeline.
- Identified systemic failures in trust, safety, and monetization.
- Designed RealMatch as a governance-first alternative.



The core issue was never dating — **it was the data systems behind it.**



DATA SOURCES AND ACKNOWLEDGMENTS

Data Sources

- App Store Reviews (Bumble, Hinge, Tinder)
- X (formerly Twitter)
- Amazon Web Services
- Figma (UX Prototyping):
<https://www.figma.com/design/jgs9WRLlfCjAirgZASRbyn/ReallMatch-Prototype?node-id=3-447&t=Eh8Pmge4wwPr8vK9-1>
- Gao, H., Yin, H., Zheng, Z., & Wang, H. (2024). Online dating apps and the association with emotional reactions: A survey on the motivations, compulsive use, and subjective online success of Chinese young adults in online dating. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 18(3), Article 3.
<https://cyberpsychology.eu/article/view/35713>
- Huang, T. F., Chou, Y. T., Chan, C. H., & Chiu, Y. C. (2023). Adolescent use of dating applications and the associations with online victimization and psychological distress. *Journal of Adolescence*. <https://pubmed.ncbi.nlm.nih.gov/37998650/>
- GitHub: https://github.com/dashduzit/swipe_right_on_data

Acknowledgments

- Thank you so much Professor Bao!