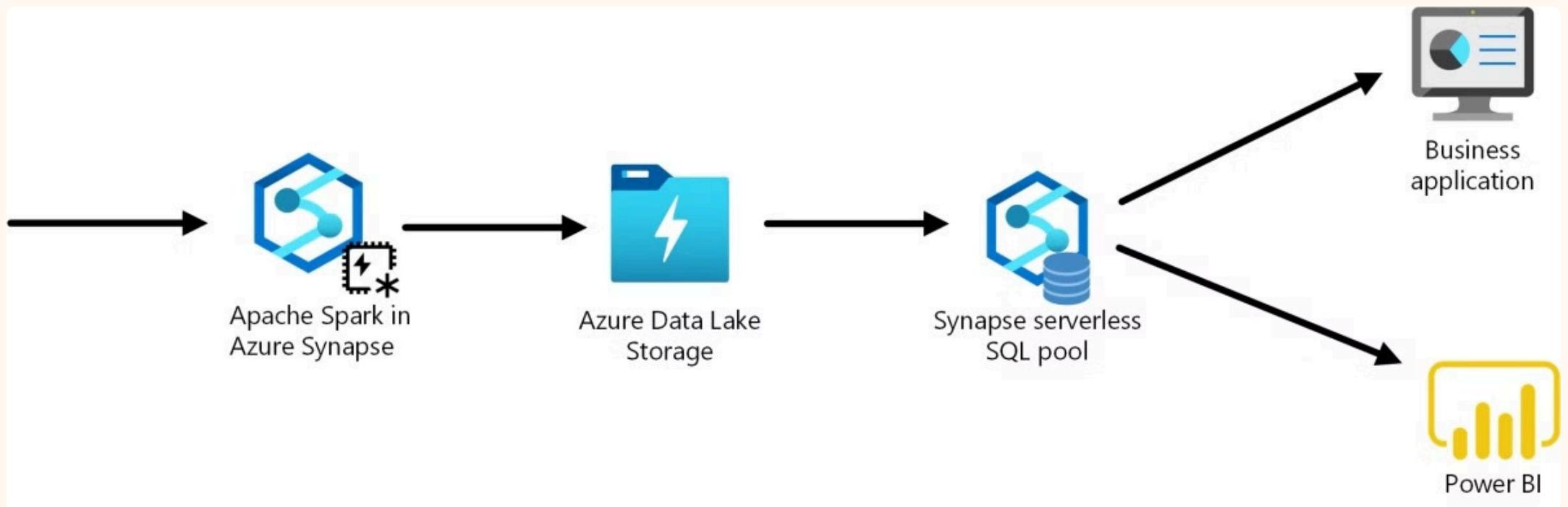


End-to-End Azure Synapse Analytics



Project Workflow

Data Extraction

Identify target websites and use automated scripts to extract relevant data.

Data Transformation

Cleanse and normalize the extracted data to ensure consistency and accuracy.

Data Storage

Securely store the cleaned data in a scalable cloud-based data lake to build Data warehouse model.

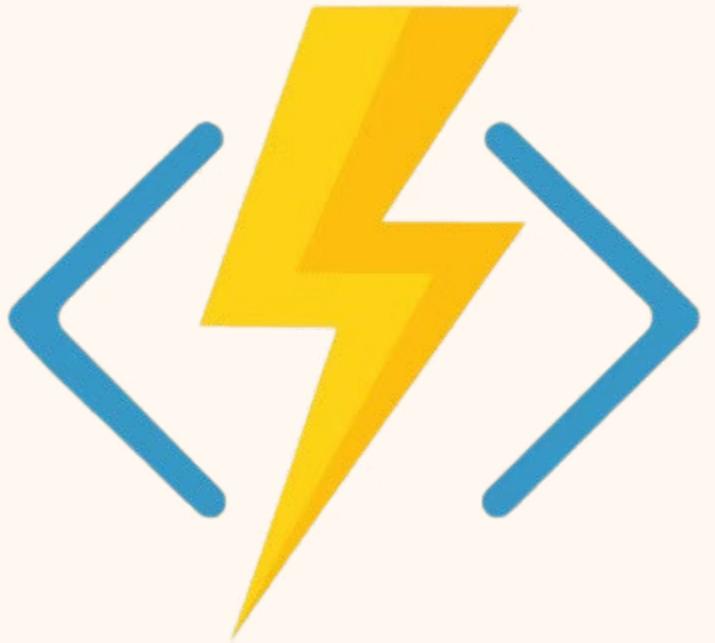
Data Analysis

Leverage advanced analytics to uncover insights and trends from the aggregated data.



Challenges in Amazon Site Scraping

- Dynamic Content
- Anti-Scraping Measures
- Data Complexity
- Legal Considerations



Leveraging Azure Function for Data Upload



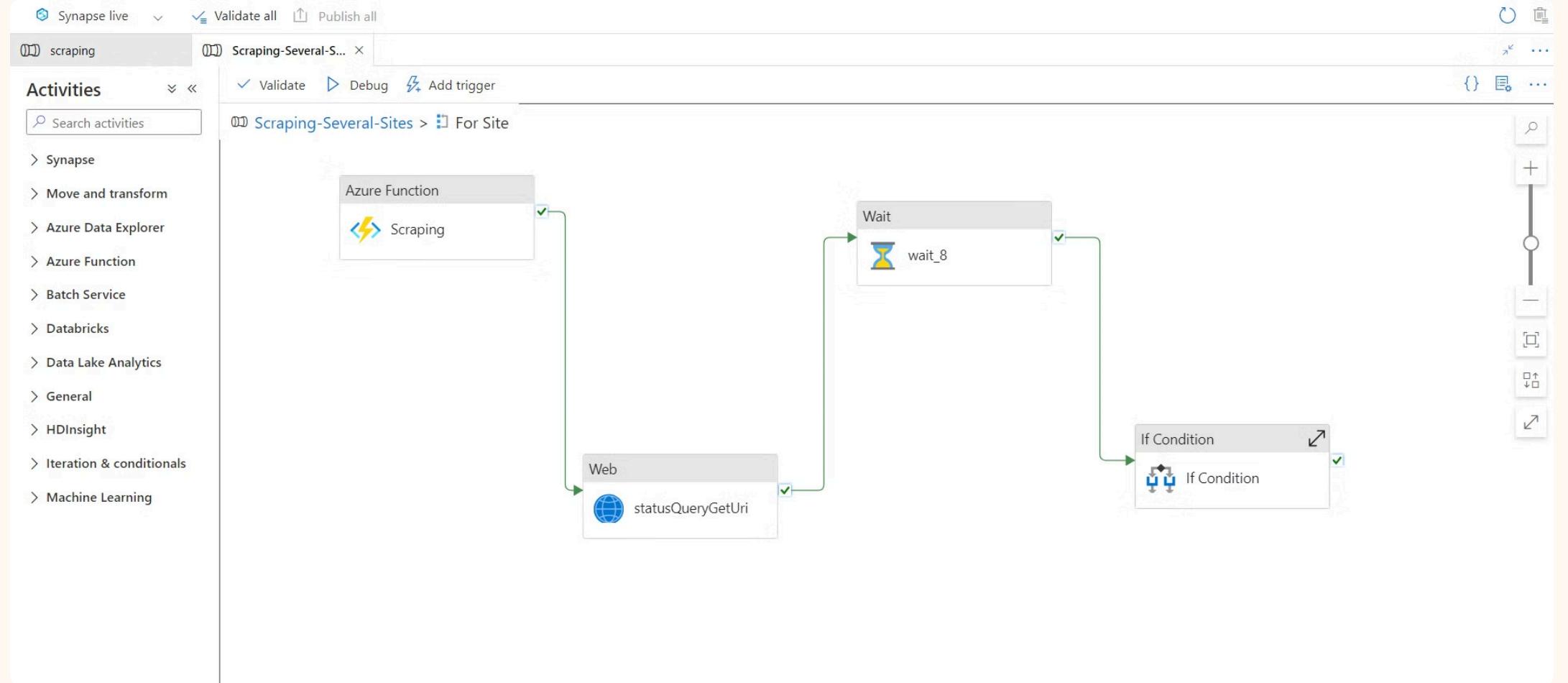
Scalable Cloud Infrastructure

Azure Function provides a flexible, serverless platform to handle high-volume data uploads.



Secure Data Storage

Extracted data is automatically stored in a scalable cloud data warehouse.



Synapse live ▾ Validate all Publish all

scraping Scraping-Several-S... X

Activities << Validate Debug Add trigger

Search activities

> Synapse
> Move and transform
> Azure Data Explorer
> Azure Function
> Batch Service
> Databricks
> Data Lake Analytics
> General
> HDInsight
> Iteration & conditionals
> Machine Learning

ForEach For Site

Parameters Variables Settings Output

Pipeline run ID: e62bd9dc-dbc1-4bbb-a4d4-35d311d4fe66 ⏪ ⏴ ⏵ Pipeline status ✓ Succeeded View debug run consumption

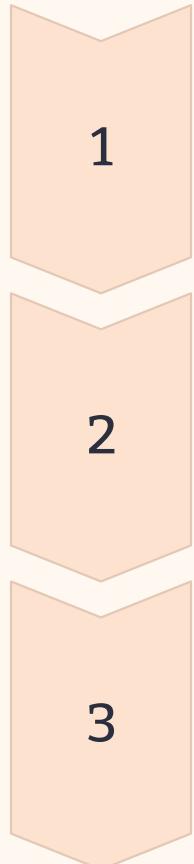
All status List

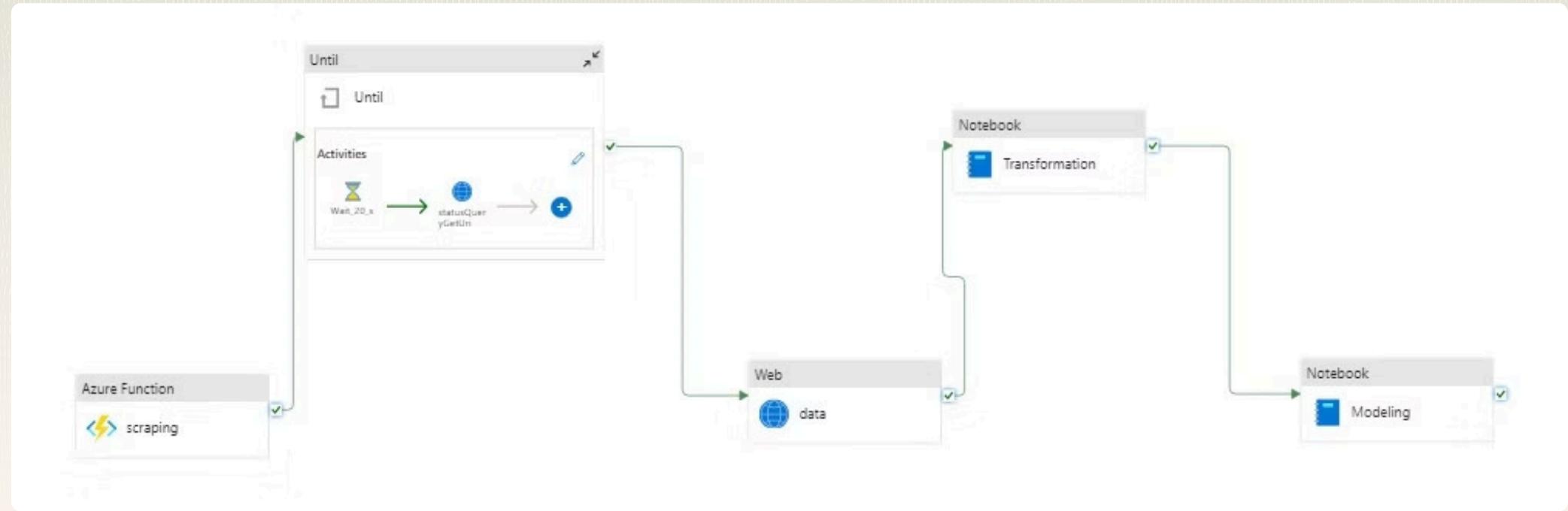
Showing 1 - 20 of 20 items

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties	Activity run ID
Data	✓ Succeeded	Web	10/17/2024, 11:54:34 P	5s	AutoResolveIntegratior		0828426a-40dc-4214-a1d3-936993
If Condition	✓ Succeeded	If Condition	10/17/2024, 11:54:34 P	7s			d833db83-b170-4504-bc75-2ee597
wait_8	✓ Succeeded	Wait	10/17/2024, 11:54:25 P	9s			e44a6b36-d026-41cc-b6e3-d87706
statusQueryGetUri	✓ Succeeded	Web	10/17/2024, 11:51:15 P	3m 9s	AutoResolveIntegratior		42de1168-874f-45e7-9a96-afcd8ad
Scraping	✓ Succeeded	Azure Function	10/17/2024, 11:51:12 P	2s	AutoResolveIntegratior		bbaedba4-540b-4899-b0fb-064e35
Data	✓ Succeeded	Web	10/17/2024, 11:51:03 P	7s	AutoResolveIntegratior		4cff1c36-b161-4e75-8e22-ffa9dfd
If Condition	✓ Succeeded	If Condition	10/17/2024, 11:51:03 P	9s			08158e1c-1ab6-4c91-985e-666e54
wait_8	✓ Succeeded	Wait	10/17/2024, 11:50:53 P	9s			0ce85b1a-1010-4346-a614-05144b

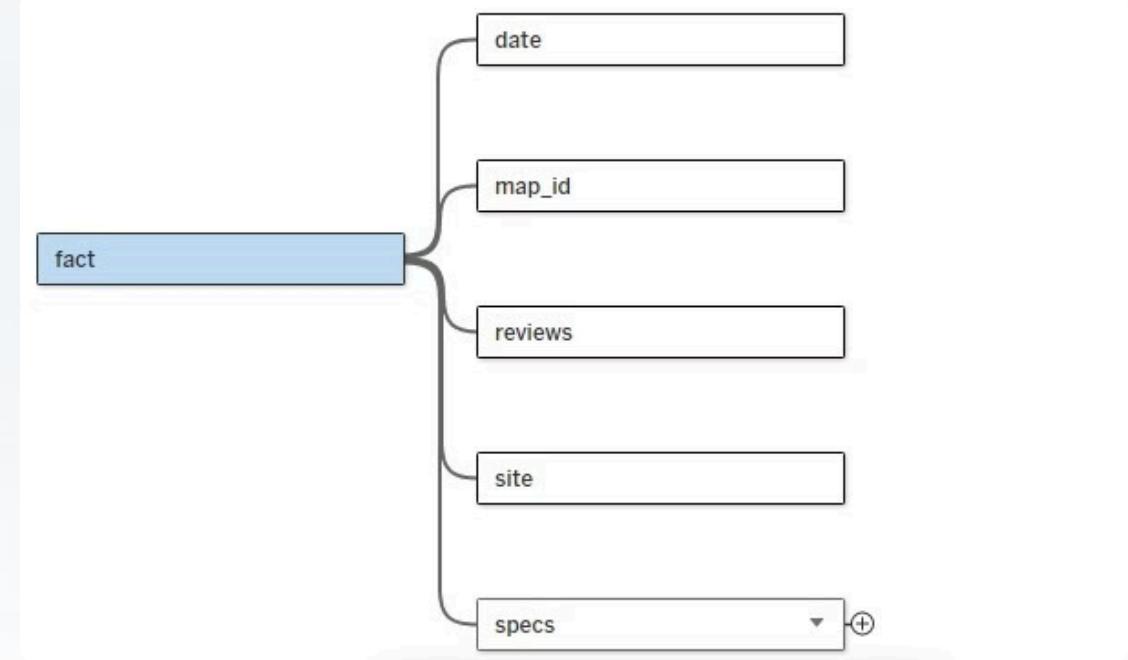
Monitor in Azure Metrics Export to CSV

Data Transformation Pipeline

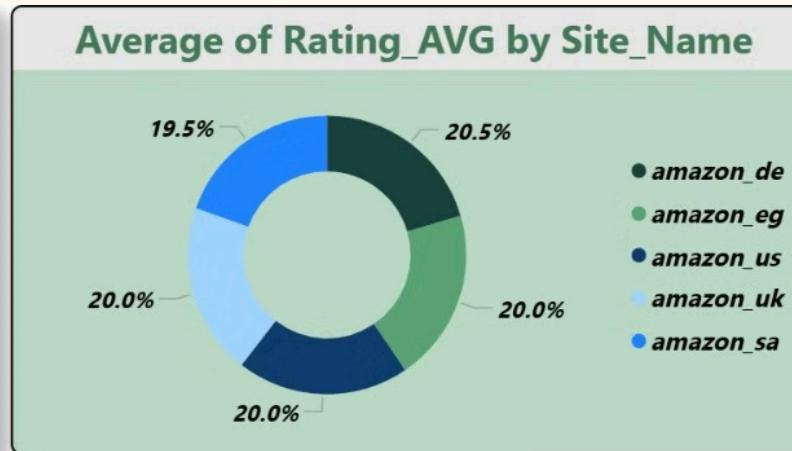
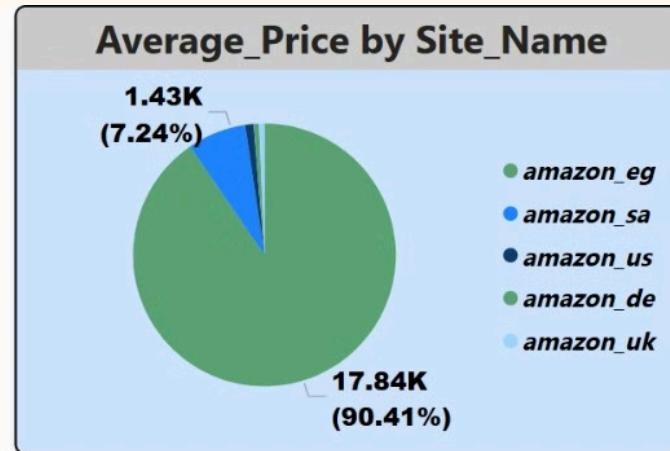




Data Warehousing



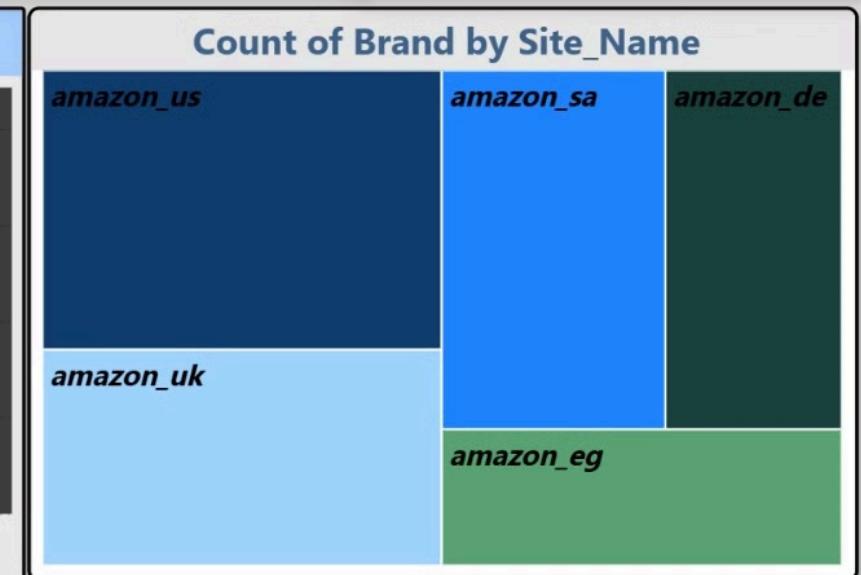
Dashboard



Brand

A list of brands displayed vertically. The brands are:

- 8849
- 8849pro
- a slice red
- accof
- acogedor
- agm
- alcatel
- annadus



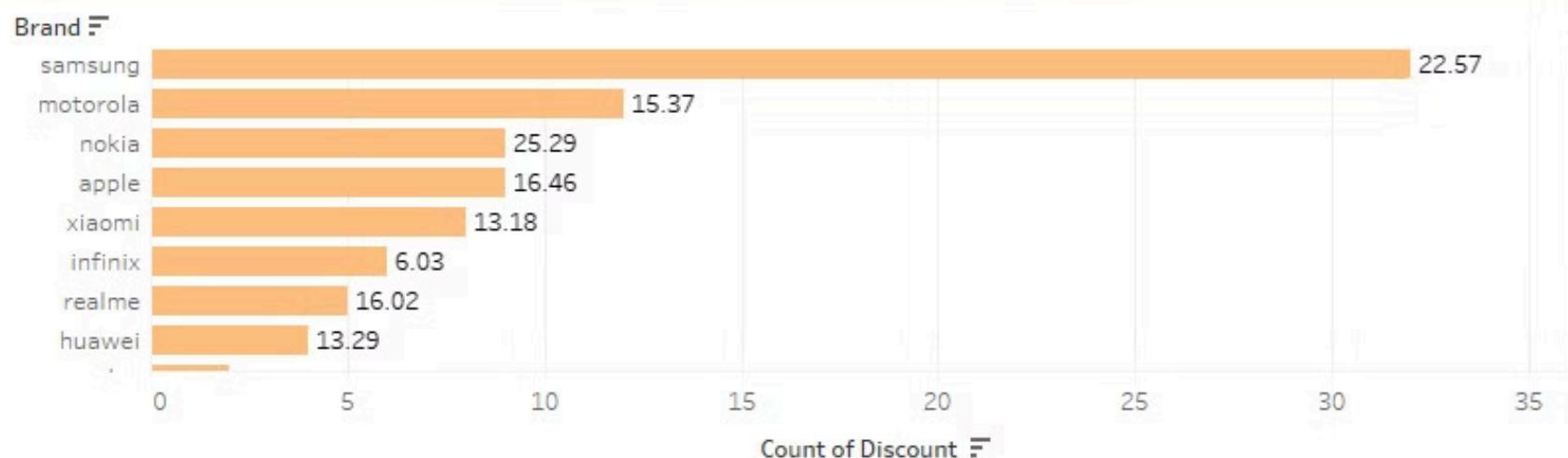
Best Discount per Brand by Site

Site Name:

 Brand Set:

Brand	Model Name	URL for Max Discount ↗	Price Before	Price	Currency	Rating Avg	Discount
apple	15	https://www.amazon.sa...	3,799	2,965	sar	4.5	21.95
		https://www.amazon.sa...	3,799	2,965	sar	4.5	21.95
		https://www.amazon.sa...	3,799	2,965	sar	4.5	21.95
huawei	huawei nova 12 se	https://www.amazon.sa...	1,149	899	sar	3.5	21.76
infinix	gt20 pro	https://www.amazon.sa...	1,099	1,017	sar	3.7	7.46
		https://www.amazon.sa...	1,099	1,017	sar	3.7	7.46
nokia	nokia ta-1114 106	https://www.amazon.sa...	88	53	sar	3.2	39.77
		https://www.amazon.sa...	88	53	sar	3.2	39.77
realme	rmx3840	https://www.amazon.sa...	1,899	1,483	sar	3.5	21.91
		https://www.amazon.sa...	699	629	sar	3.9	10.01
	rmx3997	https://www.amazon.sa...	699	629	sar	3.9	10.01
samsung	s24 ultra	https://www.amazon.sa...	5,399	3,400	sar	4.2	37.03
		https://www.amazon.sa...	5,399	3,400	sar	4.2	37.03

No. Discount For Every Brand



Best Deals Of The Day

Brand	Model Name	URL for Max Discount ↗	Price Before	Price	Currency	Rating Avg	Discount
apple	15	https://www.amazon.sa...	3,799	2,965	sar	4.5	21.95
		https://www.amazon.sa...	3,799	2,965	sar	4.5	21.95



Brand	Model_Name	Rating_AVG	Price			
blackview	tb-2024-wave 6c-blue-yq-new	5.00	9.00 84500.00			
bvna	i16 pm_240828	5.00				
bvna	i24 ultra	5.00				
catcalls	catcalls-i15pro	5.00				
doogee	blade 10	5.00				
doogee	n50pro	5.00				
fadnou	i15 pm	5.00				
hmd	skyline	5.00				
honor	honor 200 smart	5.00				
Brand	Storage	RAM_GB	Price	currency	Discount	color
apple	128.00		1982.00	sar	12%	yellow
apple	128.00		2679.00	sar		yellow
samsung	256.00	12.00	39850.00	egp		yellow
samsung	256.00	12.00	39950.00	egp	23%	yellow
samsung	256.00	12.00	51800.00	egp		yellow
geekoto			31.00	eur		white
hipipooo		1.00	23.00	eur		white
amazon_de			amazon_eg			
amazon_sa			amazon_uk			
amazon_us						

Go back

Site_Product

Tables

Data update...

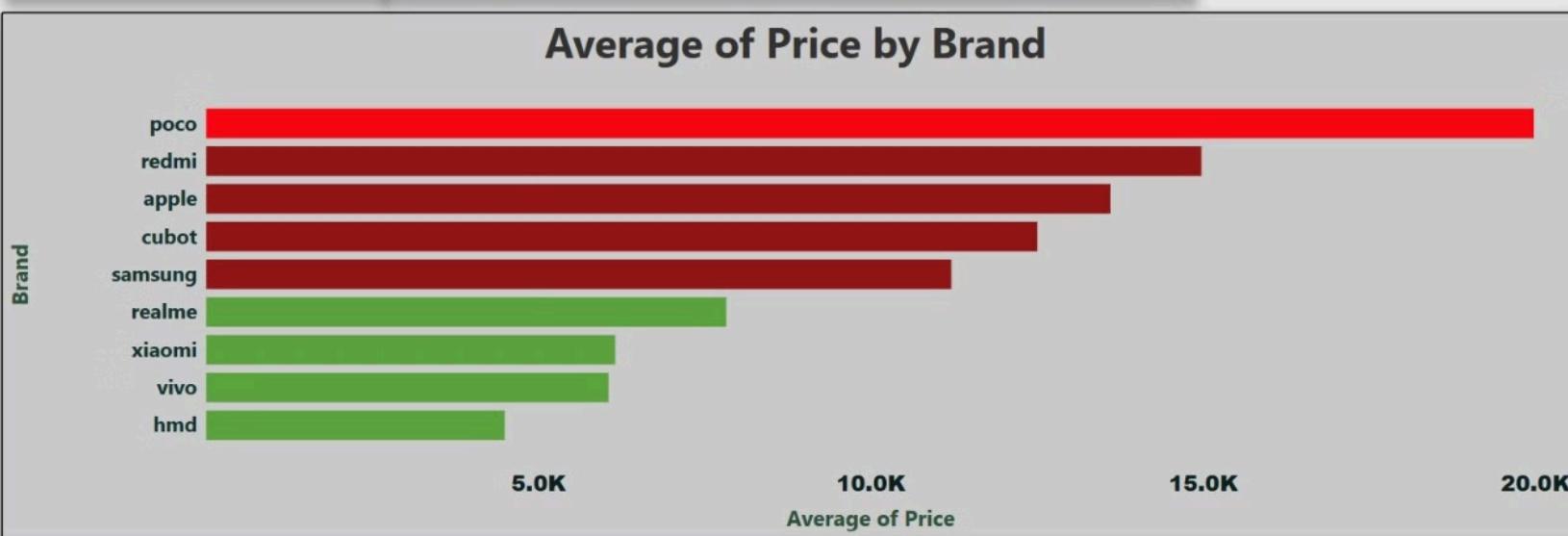
1197**Count of Brand****Currency**egp gbp usd
eur sar**Brand_name**

8849	f2ftlk	mechanicalzeal
8849pro	fadnou	mmv
a slice red	fdmdaf	moman
accof	feilei	motorola
acogedor	fossibot	mp myphone
aqm	fuxinya	msmqqa
alcatel	gabb	mtgud
annadue	geekoto	nawfait
aozveiz	generic	neoman
apexel	gesmatek	niega
apple	girasole	nokia

Rating_AVG

1.00 5.00

Filters

Average of Price by Brand

← Go back

Avg_price

Tables

Data update...



+

79%



Made with Gamma

