



5,158 ล้านคน

คิดเป็น 64.4%

ของประชากร

ทั่วโลก

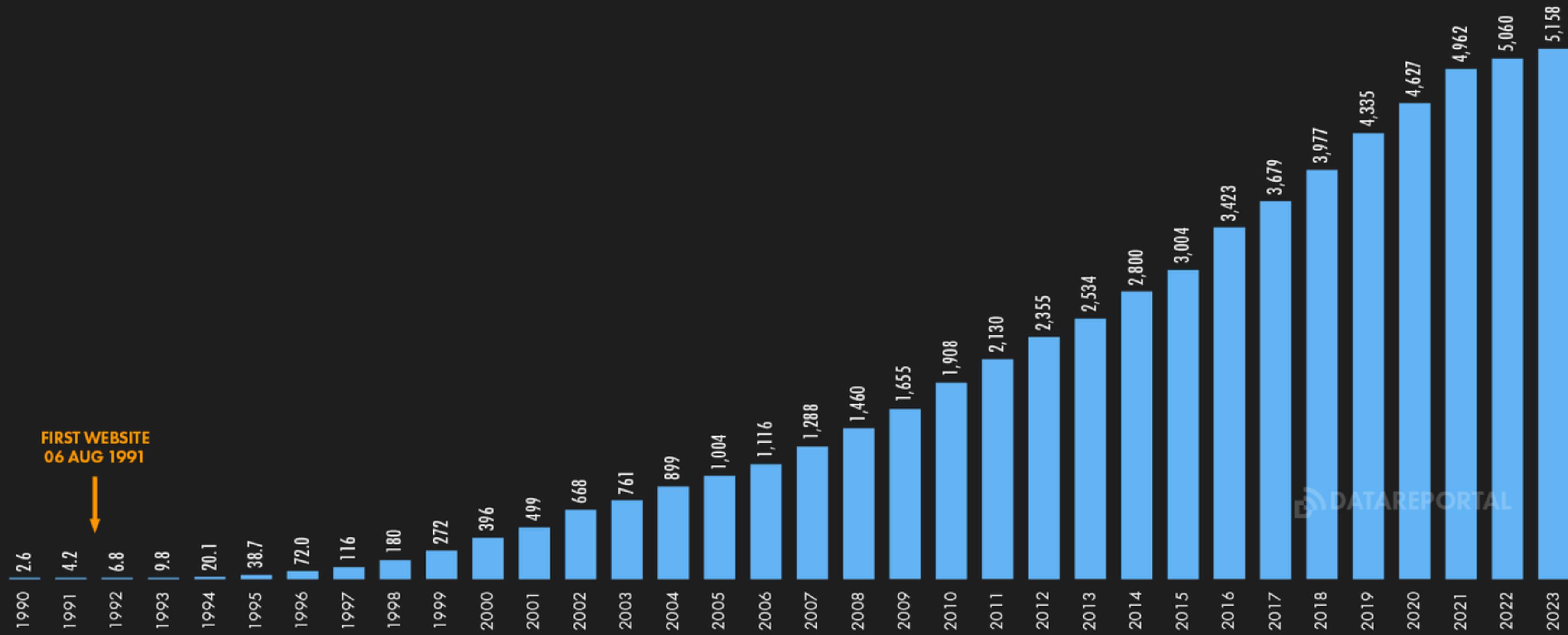
DIGITAL 2023

GLOBAL OVERVIEW REPORT

THE ESSENTIAL GUIDE TO THE WORLD'S CONNECTED BEHAVIOURS

INTERNET USERS: TIMELINE

NUMBER OF INTERNET USERS BY YEAR (IN MILLIONS)



SOURCES: KEPIOS ANALYSIS; ITU; GSMA INTELLIGENCE; EUROSTAT; WORLD BANK; GOOGLE'S ADVERTISING RESOURCES; CIA WORLD FACTBOOK; CNNIC; APJI; KANTAR & IAMAI; LOCAL GOVERNMENT AUTHORITIES; UNITED NATIONS. **NOTES:** THE TIME REQUIRED TO COLLECT, PROCESS, AND REPORT INTERNET USER RESEARCH DATA MAY MEAN THAT USER FIGURES AND GROWTH TRENDS FOR RECENT PERIODS UNDER-REPRESENT ACTUAL VALUES. SEE [NOTES ON DATA](#) FOR FURTHER DETAILS. **COMPARABILITY:** SOURCE AND BASE CHANGES. FIGURES MAY NOT MATCH OR CORRELATE WITH FIGURES PUBLISHED IN PREVIOUS REPORTS. ALL FIGURES USE THE LATEST AVAILABLE DATA, BUT SOME SOURCE DATA MAY NOT HAVE BEEN UPDATED IN THE PAST YEAR. SEE [NOTES ON DATA](#) FOR DETAILS.



BIG DATA

“ หมายความว่า ”

“ຫວຸມລົກທີ່ມີປຣິມານມາດຈັດກາຮຽກ”

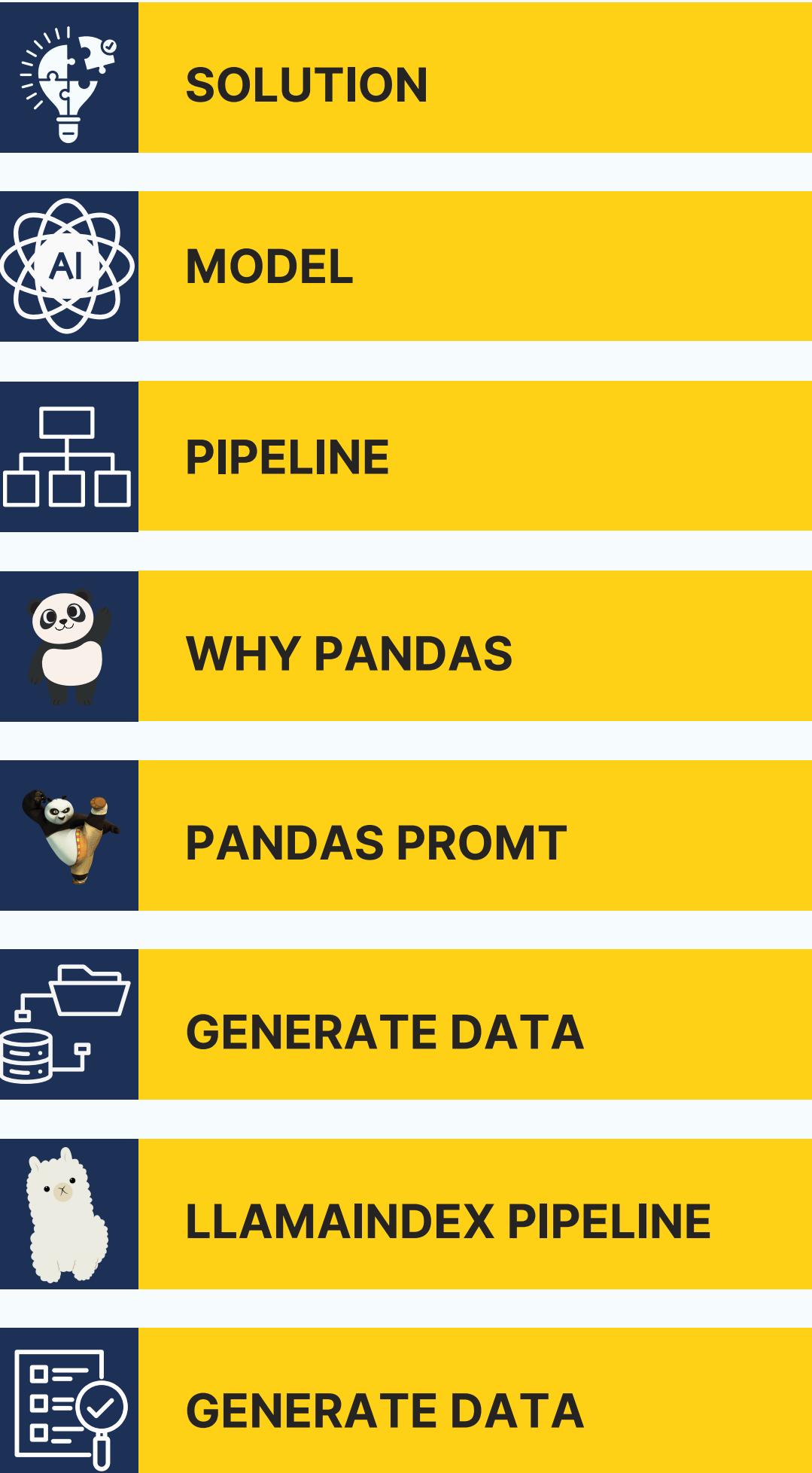
The screenshot shows the Microsoft Power Query Editor interface. A summary table is being transformed, with the formula `= Table.TransformColumnTypes` applied. The editor displays properties like Data Type: Text and columns such as RegNo, Question, CLO-1, PLO-1, Taxo, and Value. The preview pane shows the transformed data.

RegNo	Question	CLO-1	PLO-1	Taxo	Value
18-ME-1	MidQ1	CLO-1	PLO-1	C1	7
18-ME-5	MidQ2	CLO-2	PLO-2	C2	7
18-ME-1	MidQ3	CLO-2	PLO-2	C2	7
18-ME-1	MidQ4	CLO-3	PLO-3	C3	7
18-ME-1	MidQ5	CLO-3	PLO-3	C3	7
18-ME-1	MidQ6	CLO-3	PLO-3	C3	7
18-ME-1	MidQ7	CLO-3	PLO-3	C3	7
18-ME-1	FinalQ1	CLO-2	PLO-2	C3	7
18-ME-5	MidQ2	CLO-2	PLO-2	C2	7
18-ME-5	MidQ3	CLO-2	PLO-2	C2	7
18-ME-5	MidQ4	CLO-3	PLO-3	C3	7
18-ME-5	MidQ5	CLO-3	PLO-3	C3	7
18-ME-5	MidQ6	CLO-3	PLO-3	C3	7
18-ME-5	MidQ7	CLO-4	PLO-3	C3	7
18-ME-5	FinalQ1	CLO-2	PLO-2	C3	7
18-ME-10	MidQ1	CLO-1	PLO-1	C1	7
18-ME-10	MidQ2	CLO-2	PLO-2	C2	7
18-ME-10	MidQ3	CLO-2	PLO-2	C2	7
18-ME-10	MidQ4	CLO-3	PLO-3	C3	7
18-ME-10	MidQ5	CLO-3	PLO-3	C3	7
18-ME-10	MidQ6	CLO-3	PLO-3	C3	7
18-ME-10	MidQ7	CLO-4	PLO-3	C3	7
18-ME-10	FinalQ1	CLO-2	PLO-2	C3	7
18-ME-20	MidQ1	CLO-1	PLO-1	C1	7
18-ME-20	MidQ2	CLO-2	PLO-2	C2	7
18-ME-20	MidQ3	CLO-2	PLO-2	C2	7
18-ME-20	MidQ4	CLO-3	PLO-3	C3	7
18-ME-20	MidQ5	CLO-3	PLO-3	C3	7
18-ME-20	MidQ6	CLO-3	PLO-3	C3	7
18-ME-20	MidQ7	CLO-4	PLO-3	C3	7
18-ME-20	FinalQ1	CLO-2	PLO-2	C3	7
18-ME-27	MidQ1	CLO-1	PLO-1	C1	7
18-ME-32	MidQ2	CLO-2	PLO-2	C2	7
18-ME-32	MidQ3	CLO-2	PLO-2	C2	7
18-ME-41	MidQ4	CLO-3	PLO-3	C3	7
18-ME-41	MidQ5	CLO-3	PLO-3	C3	7
18-ME-44	MidQ6	CLO-3	PLO-3	C3	7
18-ME-44	MidQ7	CLO-4	PLO-3	C3	7
18-ME-47	FinalQ1	CLO-2	PLO-2	C3	7
18-ME-5	MidQ1	CLO-1	PLO-1	C1	7
18-ME-53	MidQ2	CLO-2	PLO-2	C2	7
18-ME-53	MidQ3	CLO-2	PLO-2	C2	7
18-ME-56	MidQ4	CLO-3	PLO-3	C3	7
18-ME-56	MidQ5	CLO-3	PLO-3	C3	7
18-ME-63	MidQ6	CLO-4	PLO-3	C3	7
18-ME-63	MidQ7	CLO-2	PLO-2	C3	7
18-ME-70	FinalQ1	CLO-1	PLO-1	C1	7

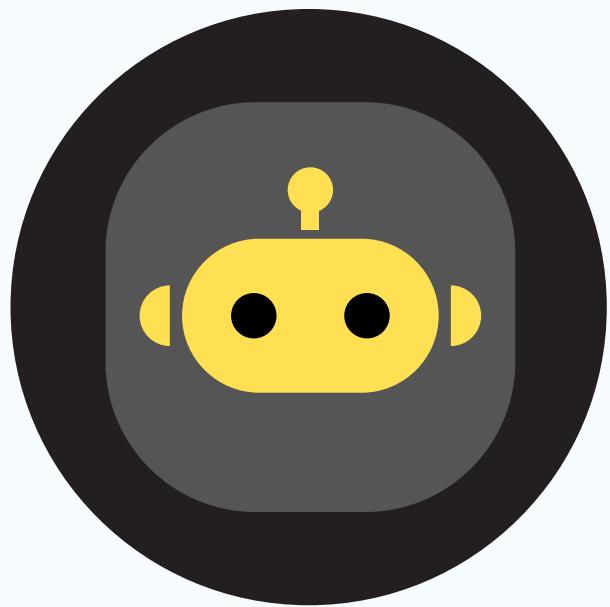
	amount	Employee	Department	Expense Category	Expense Type	LEN(Expense Type)
1	50	Derek Stand	Sales	500 Petrol	500	10
2	130	Derek Stand	Sales	300 Hotel	300	9
3	43.16	Derek Stand	Sales	100 Food	100	8
4	12.45	Derek Stand	Sales	100 Food	100	8
5	10.5	Derek Stand	Sales	100 Food	100	8
6	40	Mary Wells	Sales	500 Petrol	500	10
7	210	Mary Wells	Sales	400 Hotel	400	9
8	39	Mary Wells	Sales	100 Food	100	8
9	12.45	Mary Wells	Sales	100 Food	100	8
10	11.12	Mary Wells	Sales	100 Sundries	100	12
11	5	Mary Wells	Sales	600 Stationary	600	14
12	45	Paul Simmons	Support	500 Train	500	9
13	130	Paul Simmons	Support	300 Hotel	300	9
14	43.16	Paul Simmons	Support	100 Food	100	8
15	12.45	Paul Simmons	Support	100 Food	100	8
16	10.5	Paul Simmons	Support	100 Food	100	8
17	2.5	Paul Simmons	Support	500 Taxi	500	8



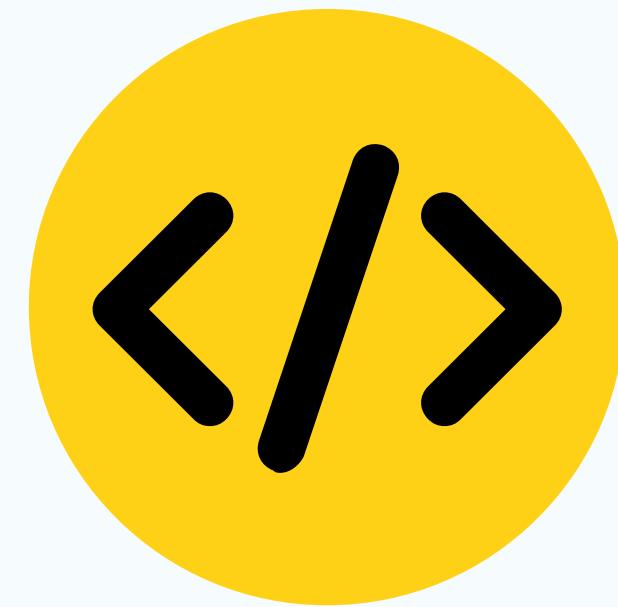
OVERVIEW



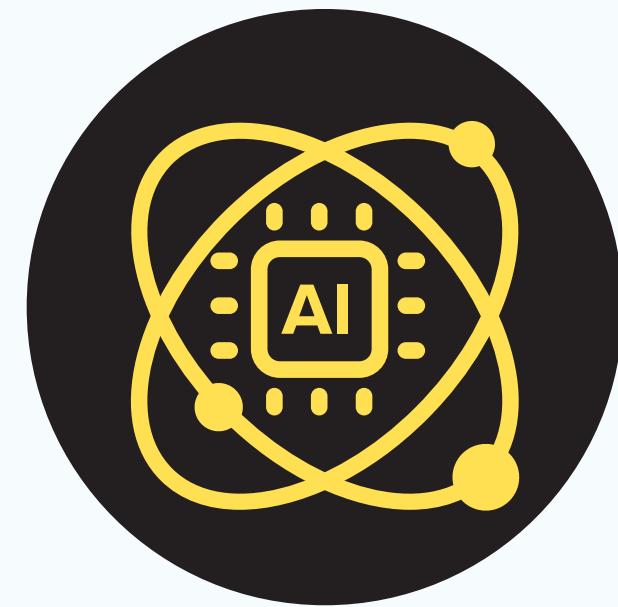
SOLUTION



OpenthaiGPT 13b
+
llamaindex

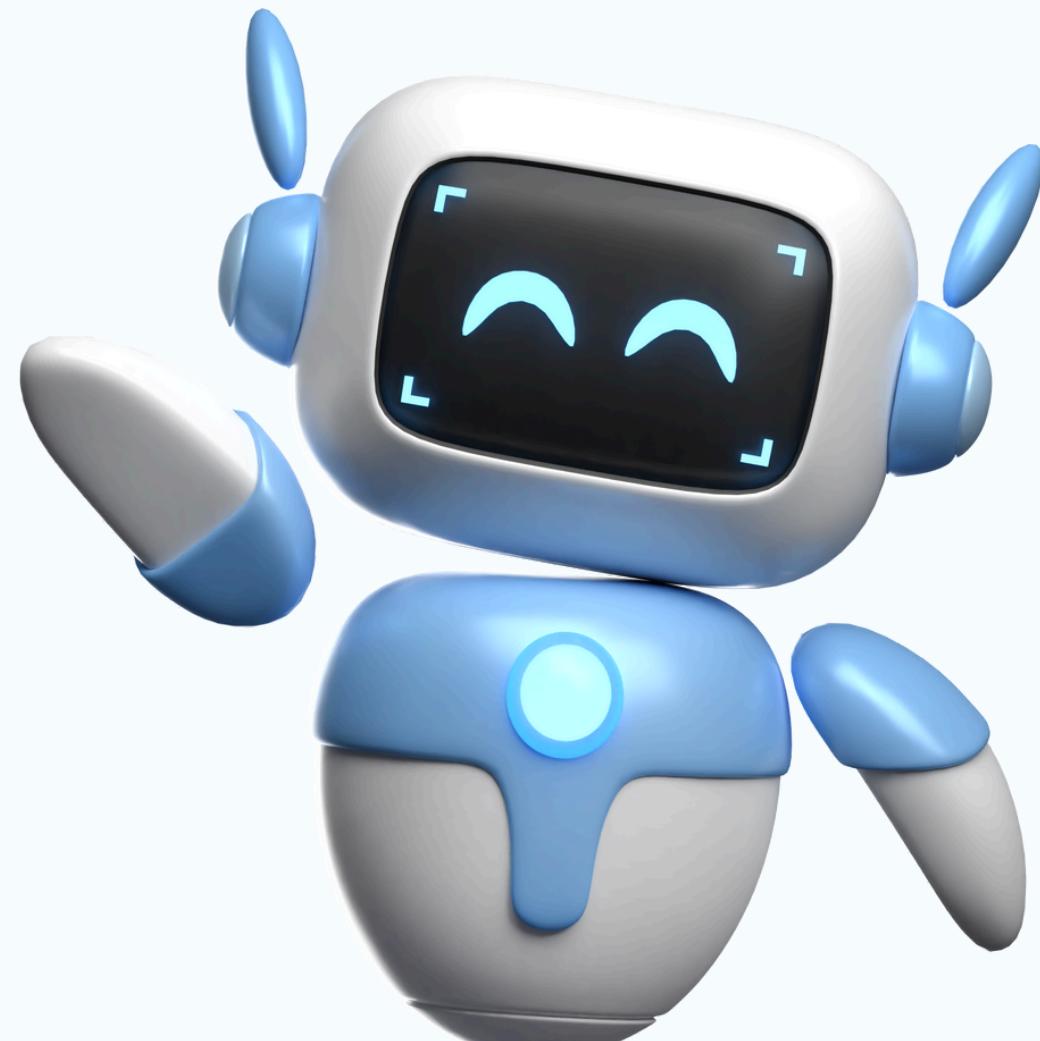


Query data from
data base



Generate Answers
based on the data

MODEL



OpenthaiGPT 13b



Base Model: LLaMa V2

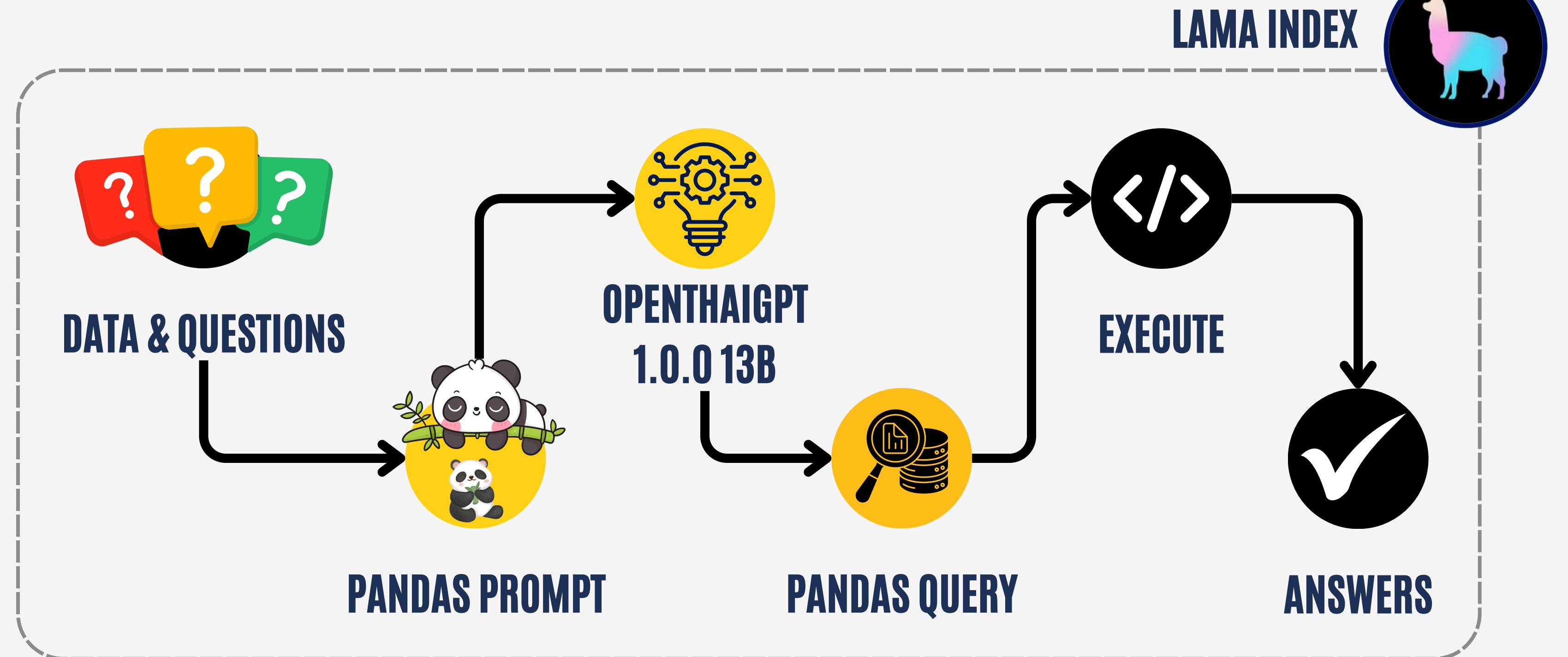


Training Data: 15B Pretrained
Thai tokens



Fine Tune: 10,000 Thai words

PIPELINE



PROJECT

WHY PANDAS

```
SELECT SUM(AAPL_gross_profit) - SUM(AMZN_gross_profit) AS profit_difference
FROM (
    SELECT Year, Company, Gross Profit AS AAPL_gross_profit
    FROM financial_data
    WHERE Company = 'AAPL' AND Year >= 2019
) AS aapl_data
JOIN (
    SELECT Year, Company, Gross Profit AS AMZN_gross_profit
    FROM financial_data
    WHERE Company = 'AMZN' AND Year >= 2019
) AS amzn_data
ON aapl_data.Year = amzn_data.Year;
```

SQL



“A single line of Pandas Query code achieves the same result as complex SQL queries, improving clarity and maintainability.”

Pandas

```
profit_difference = (df[(df['Company'] == 'AAPL') \
& (df['Year'] >= 2019)]['Gross Profit'].sum()) - \
(df[(df['Company'] == 'AMZN')\
& (df['Year'] >= 2019)]['Gross Profit'].sum())
```

PROJECT

PANDAS PROMT

```
pandas_prompt_str = (
    "You are working with a pandas dataframe in Python.\n"
    "The name of the dataframe is df.\n"
    "This is the example of table:\n"
    "{df_str}\n\n"
    "ตารางมีคอลัมน์ดังนี้:\n"
    "{df_col}\n\n"
    "Follow these instructions:\n"
    "{instruction_str}\n"
    "Query: {query_str}\n\n"
    "Expression:"
)
```



**“Prompt with Details
for decreased hallucination”**

```
instruction_str = (
    "1. Convert the query to executable Python code using Pandas.\n"
    "2. The final line of code should be a Python expression that can be called with the eval()\nfunction.\n"
    "3. The code should represent a solution to the query.\n"
    "4. PRINT ONLY THE EXPRESSION.\n"
    "5. Do not quote the expression.\n"
)
```

PIPELINE

LAMA INDEX



IMPROVE DATA QUALITY

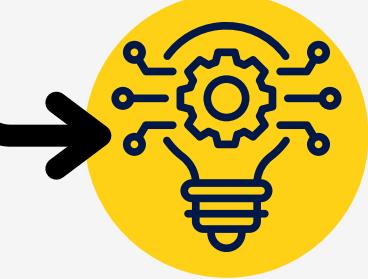


DATA & QUESTIONS



PANDAS PROMPT

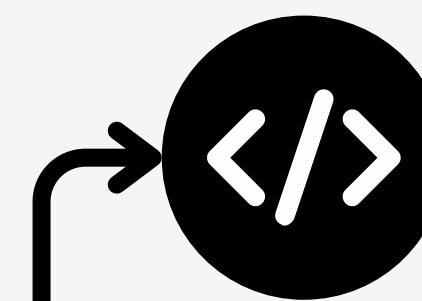
OPENTHAIGPT
1.0.0 13B



PANDAS QUERY



EXECUTE



ANSWERS





AUTOMATIC GENERATION WITH

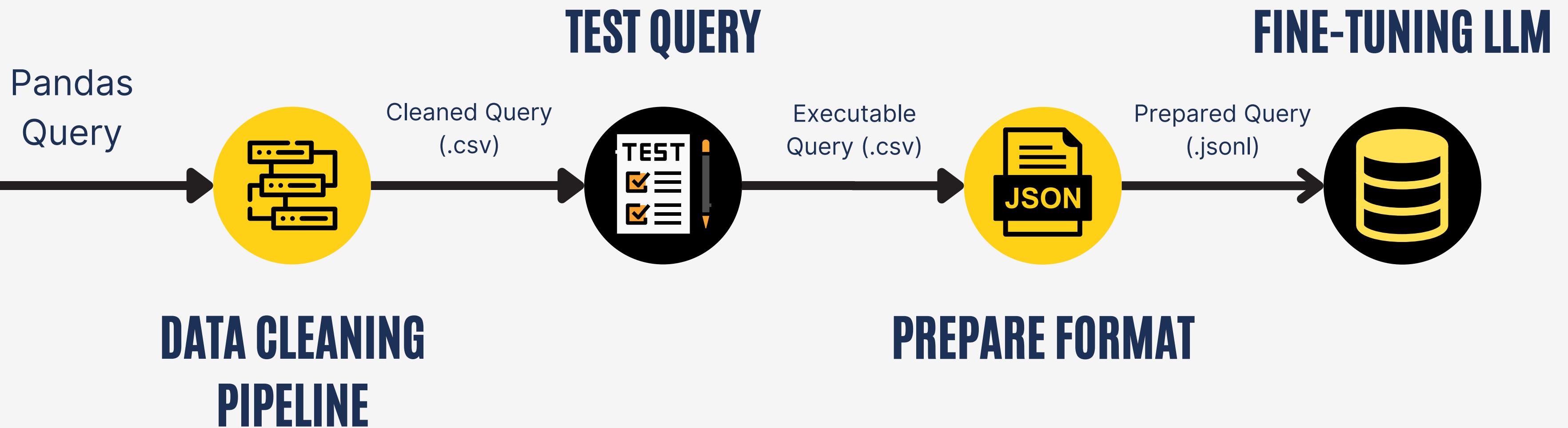
Gemini API





AUTOMATIC GENERATION WITH

Gemini API



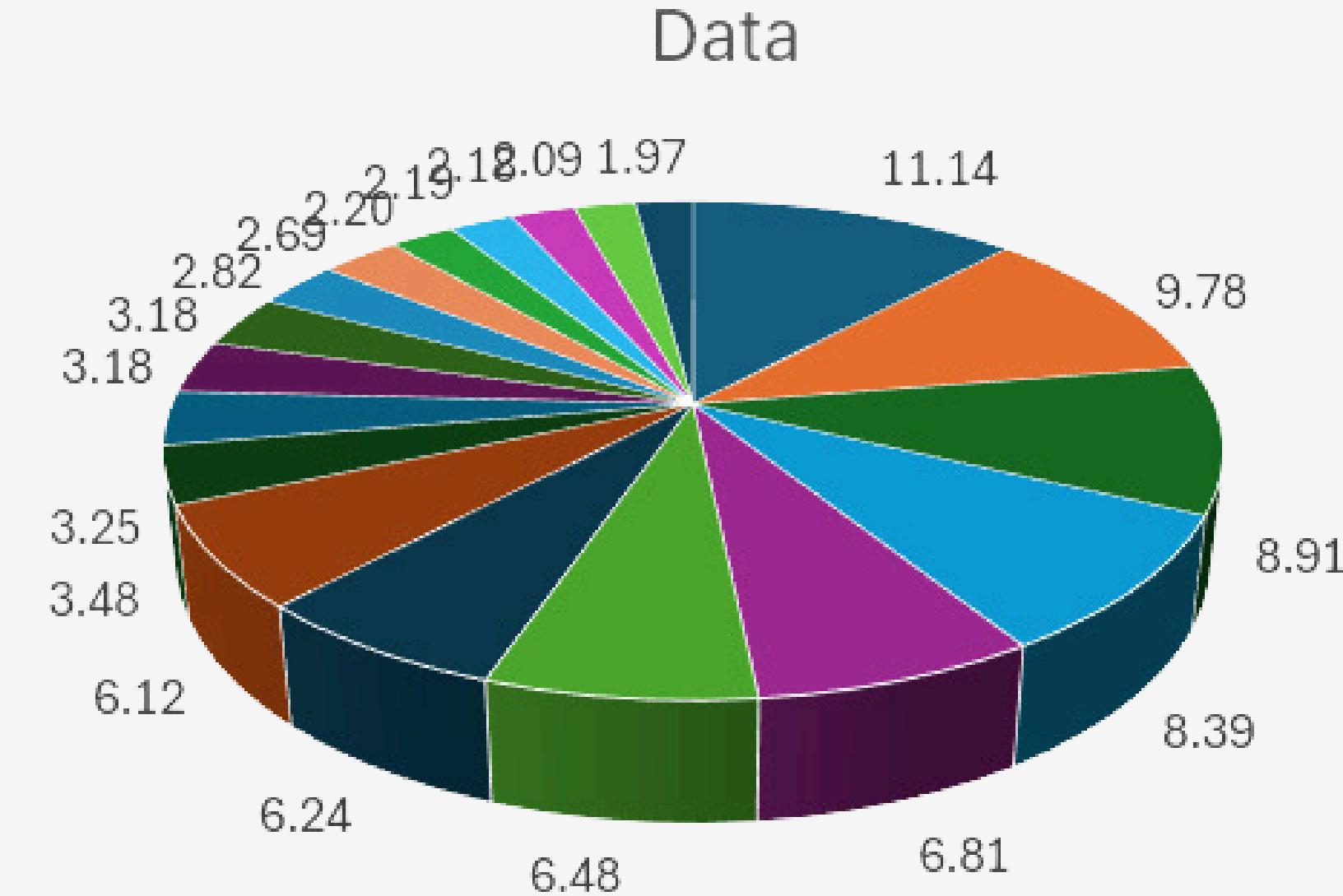
Data

Generate dataset

- Training
- Before clean 28,891 rows
- After clean 24,376 rows

- Table 2:Online shopping
- Car price prediction
- Consumer Behavior and shopping habits
- Loan eligible
- transaction data3
- Stock market data
- Bkie-share usage dataset

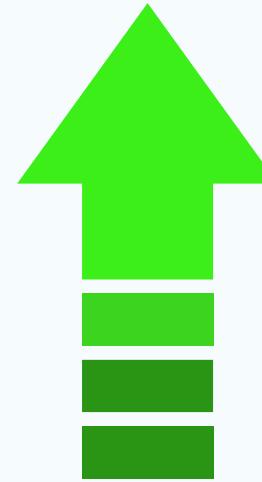
- Table C : Spotify
- Table1 :Financial data
- Table D : Customer_support_tickets
- Employee Salaries analysis
- Dairy goods sales
- coffee shop sales analisis
- Bank customer





RESULT

PUBLIC DATASET



53 %

ITERATION

\approx 2.4 S/REQ

THANK YOU!

Dataset



We are Kiddee

<https://huggingface.co/datasets/AIA/T/Kiddee-data1234>

Model



<https://huggingface.co/AIAT/Kiddee-qatable1>