

# Project Predict the Future

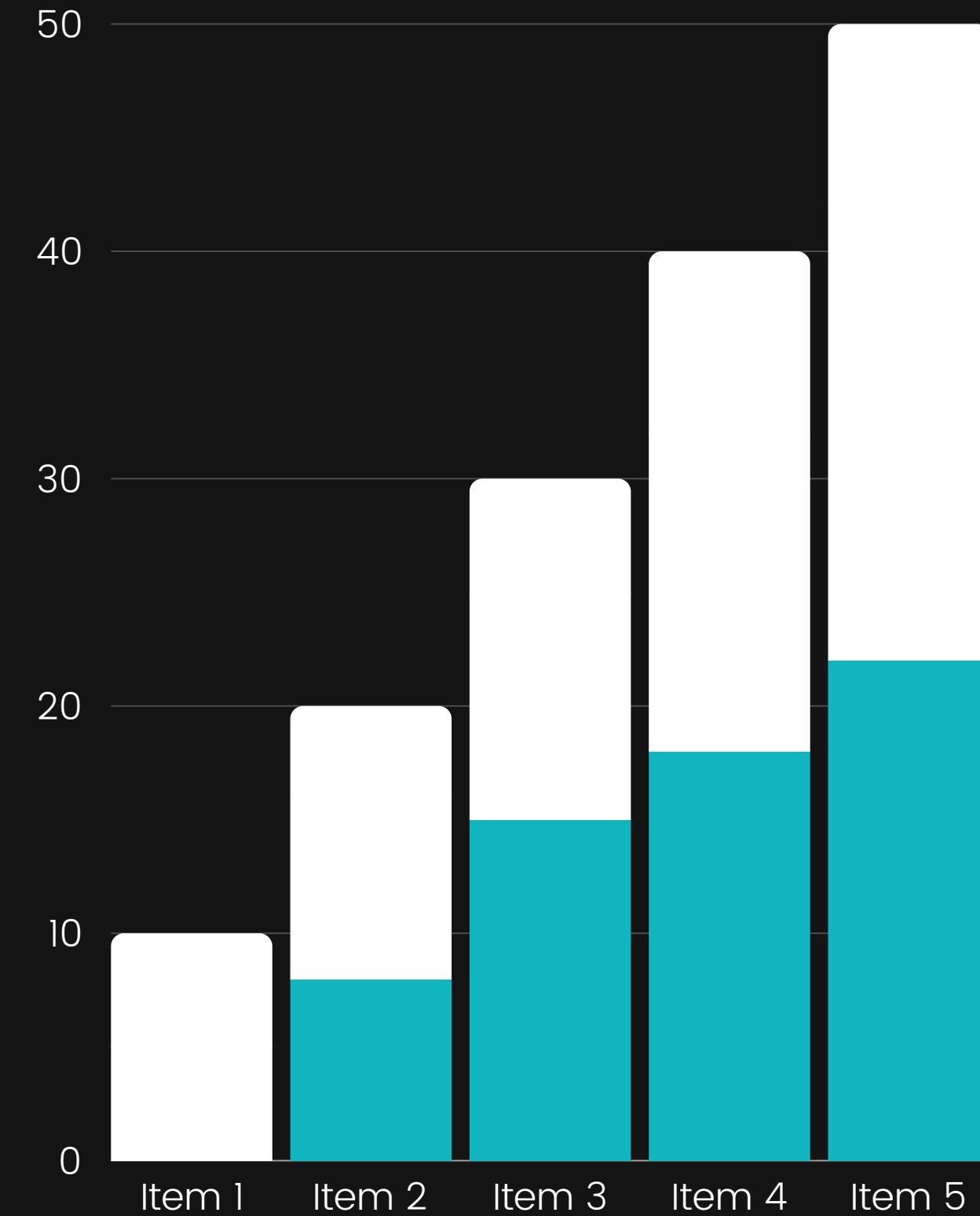
Instructor: Thai Thi Thanh Thao

Implementer: Doan Nguyen Ha My



# Current situations

Currently, big data projects predicting future products are attracting great attention from the research and business communities. Using technologies such as machine learning and predictive analytics, researchers and data experts have conducted big data analysis to predict market trends and consumer behavior.



# Objectives

Big Data and Data Analysis

---

Python Programming Language

---



# Aim

Trend prediction

---

Optimize business decisions

---

Risk mitigation

---

Increase chances of success

---

# Primary Research

## PRIMARY RESEARCH SURVEY DATA

Primary research survey data is data collected from surveys conducted directly with individuals or organizations to gather information and opinions on a specific topic.

## PRIMARY RESEARCH INTERVIEW

Primary research interview is a method of collecting data directly from participants through conducting personal interviews.

## PRIMARY RESEARCH OBSERVATION

Primary research observation is a method of data collection in research in which researchers directly observe and record specific behaviors, events, or phenomena of interest.

## PRIMARY RESEARCH DOCUMENTS

Primary research documents are a method of data collection in research in which the researcher uses original or previously unreviewed documents to gather information and data.

# Secondary Research

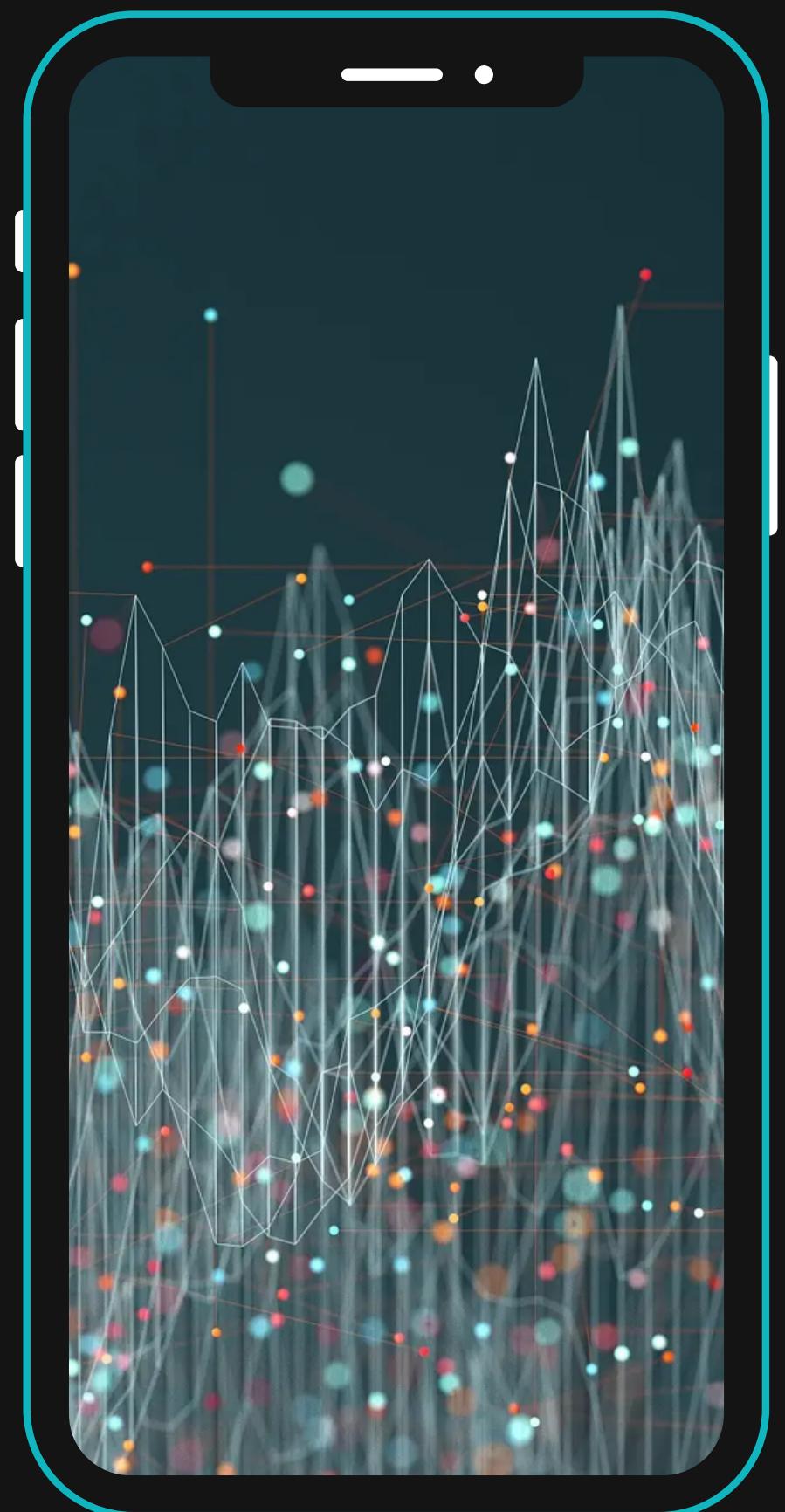
Secondary Research, or secondary research, is the process of collecting and analyzing information from existing data sources, instead of creating new data through methods such as direct surveys.



# Research methods

Use data from Tableau.

Use the Python language and libraries such as Pandas, NumPy, and Scikit-learn



# This implementation method has the following advantages and disadvantages

Diverse and complete data

Limits of the model

The data has been cleaned

Depends on data quality

Prediction accuracy

Resource and knowledge requirements

# Dataset

---

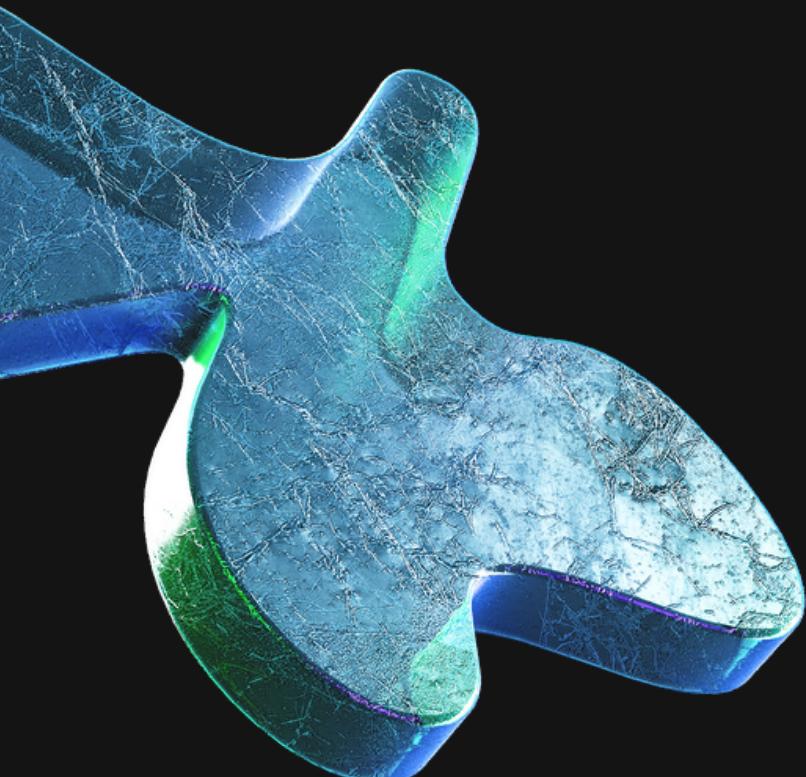
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit
2	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-1000	Furniture	Bookcases	Bush Somers	261.96	2	0	42.39
3	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-1000	Furniture	Chairs	Hon Deluxe F	731.94	3	0	146.39
4	3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	DV-13045	Darrin Van Ho	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-1000	Office Supplies	Labels	Self-Adhesive	14.62	2	0	2.92
5	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-1000	Furniture	Tables	Bretford CR4!	957.5775	5	0.45	-38.88
6	5	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-1000	Office Supplies	Storage	Eldon Fold 'N	22.368	2	0.2	4.47
7	6	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-1000	Furniture	Furnishings	Eldon Express	48.86	7	0	1.11
8	7	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-1000	Office Supplies	Art	Newell 322	7.28	4	0	1.82
9	8	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-1000	Technology	Phones	Mitel 5320 IP	907.152	6	0.2	18.14
10	9	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-1000	Office Supplies	Binders	DXL Angle-Vi	18.504	3	0.2	3.70
11	10	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-1000	Office Supplies	Appliances	Belkin F5C20	114.9	5	0	2.29
12	11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-1000	Furniture	Tables	Chromcraft R	1706.184	9	0.2	341.28
13	12	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoff	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-1000	Technology	Phones	Konftel 250 C	911.424	4	0.2	18.28
14	13	CA-2017-114412	4/15/2017	4/20/2017	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina	28027	South	OFF-PA-1000	Office Supplies	Paper	Xerox 1967	15.552	3	0.2	3.11
15	14	CA-2016-161389	12/5/2016	12/10/2016	Standard Class	IM-15070	Irene Maddo	Consumer	United States	Seattle	Washington	98103	West	OFF-BI-1000	Office Supplies	Binders	Fellowes PB2	407.976	3	0.2	13.19
16	15	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawla	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-1000	Office Supplies	Appliances	Holmes Repla	68.81	5	0.8	-1.6
17	16	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawla	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-1000	Office Supplies	Binders	Storex DuraT	2.544	3	0.8	0.73
18	17	CA-2014-105893	11/11/2014	11/18/2014	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wisconsin	53711	Central	OFF-ST-1000	Office Supplies	Storage	Stur-D-Stor Sh	665.88	6	0	13.34
19	18	CA-2014-167164	5/13/2014	5/15/2014	Second Class	AG-10270	Alejandro Gr	Consumer	United States	West Jordan	Utah	84084	West	OFF-ST-1000	Office Supplies	Storage	Fellowes Sup	55.5	2	0	1.11
20	19	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California	94109	West	OFF-AR-1000	Office Supplies	Art	Newell 341	8.56	2	0	1.71
21	20	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California	94109	West	TEC-PH-1000	Technology	Phones	Cisco SPA 50	213.48	3	0.2	4.27
22	21	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Dor	Consumer	United States	San Francisco	California	94109	West	OFF-BI-1000	Office Supplies	Binders	Wilson Jones	22.72	4	0.2	4.54

# Experiences



# Limitations

Difficult to find data, short research time, not yet implemented user interface.





# References

- [1] TopDev Blog (2019). Big Data là gì? Tất tần tật về Big Data. [online] TopDev. Available at: <https://topdev.vn/blog/big-data/> [Accessed 19 Mar. 2024].
- [2] Iboss.vn. (2022). Những điều cần biết về Big Data và dự đoán xu hướng trong tương lai – IBOSS VIETNAM. [online] Available at: <https://iboss.vn/nhung-dieu-can-biet-ve-big-data-va-du-doan-xu-huong-trong-tuong-lai/> [Accessed 19 Mar. 2024].
- [3] McBroom, S. (2022). The Best Ways for Project Managers to Predict Project Success. [online] Traxidy. Available at: <https://traxidy.com/the-best-ways-for-project-managers-to-predict-project-success/> [Accessed 19 Mar. 2024].

