

Class period 18

บทที่ 7 การแสดงผลการวิจัยเชิงข้อมูล (ต่อ)
 Visualize_Data_Comparison
 Bar chart - Grouped Bar chart

1

โหลดข้อมูล

- โฉมศึกษาจากข้อมูล search_request.csv จาก
- https://storage.googleapis.com/gcp-public-data-gcsfs/docs/590751VSCV9yCQZ379icid-4w8F37abG5C9k33-3R96ZL686w7795G7818u_457D933F31916859u68
- ดึงข้อมูลจาก google drive แล้วเชื่อม google drive โหลดไฟล์ใน colab
- from google.colab import drive
- import pandas as pd
- import os
- from datetime import datetime as dt
- from datetime import time

2

เตรียมและแปลงข้อมูลบอกเวลาให้เป็นตัวแปรชนิด timestamp

```
drive.mount('/content/drive')
path = '/content/drive/My Drive/datavis_2024_data'
data = pd.read_csv(os.path.join(path, 'search_request.csv'))
data.head()
data['search_timestamp'] =
pd.to_datetime(data['search_timestamp'], format='%Y-%m-%d
%H:%M:%S', errors='coerce')
```

3

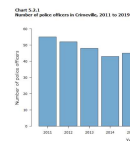
Quiz

- หน้ากราฟแสดงเปรียบเทียบปริมาณ คนเข้าถึง web Wongnai.com เป็นตัวทำนายจาก โมเดลหรือไม่

4

Bar chart (กราฟแท่ง)

- กราฟที่ใช้ในการเปรียบเทียบข้อมูลจากข้อมูล
- ส่วนประกอบของกราฟแท่ง
- ตัวกราฟแท่ง (xlabel)
- ตัวแกนกราฟแท่ง (y)
- ชื่อแกน (tick_label)
- ชื่อแกน x (plt.xlabel)
- ชื่อแกน y (plt.ylabel)



5

การแสดงตัวอักษรภาษาไทยในกราฟ matplotlib

- วิธีที่ 1
- <https://github.com/Phonopolis/sarabun-webfont/blob/master/Fonts/ThaiSarabunNew-webfont.ttf>
- https://colab.research.google.com/drive/1d7ZuGm5knc8Dh_C9t4t4yC4725Gf8o01To-8K950u55-7u8
- `fig = plt.figure()`
- `fig.savefig('font/ThaiSarabunNew-webfont.ttf')`
- `import matplotlib as mpl`
- `mpl.font_manager.fontManager.addfont('ThaiSarabunNew-webfont.ttf')`
- `mpl.rc('font', family='TH Sarabun New')`

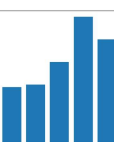
6

การสร้าง bar chart

- from matplotlib import pyplot as plt
- `plt.bar([1,2,3,4,5,6,7], [data[data['search_timestamp']].dt.dayofweek == 0].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 1].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 2].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 3].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 4].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 5].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 6].shape[0]`
- `)`

7

ผลลัพธ์ การสร้าง bar chart

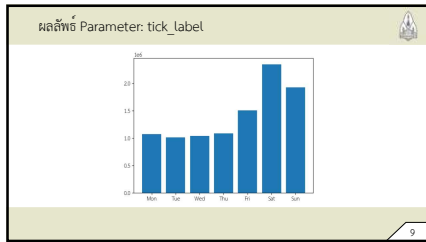


8

Parameter: tick_label (ชื่อแท่ง)

- ชื่อแท่งสามารถกำหนดโดยการใช้ Parameter tick_label (ใช้ชื่อแท่งที่ต้องการแสดง)
- `plt.bar([1,2,3,4,5,6,7], [data[data['search_timestamp']].dt.dayofweek == 0].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 1].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 2].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 3].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 4].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 5].shape[0],`
- `data[data['search_timestamp']].dt.dayofweek == 6].shape[0]`
- `), tick_labels=['Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat', 'Sun'])`

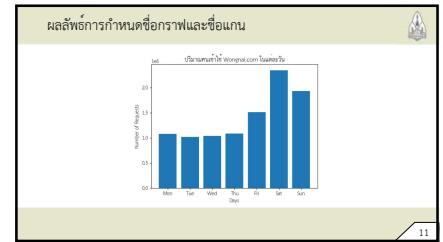
9



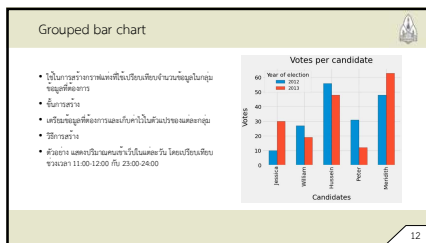
10



11



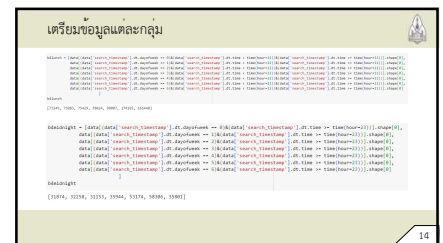
12



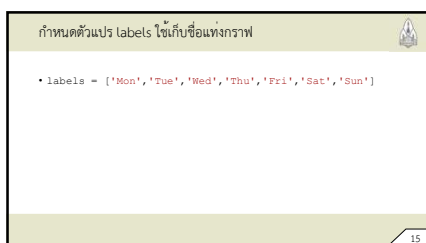
13



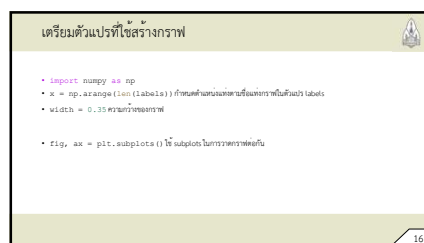
14



15



16



17

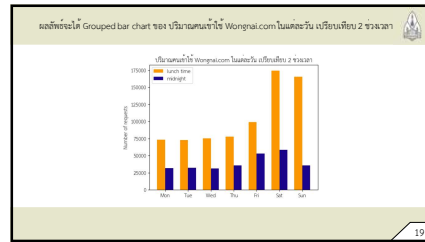


18

กำหนดส่วนประกอบกราฟต่างๆ

- `ax.set_ylabel("Number of requests")`
- `ax.set_title("ปริมาณคนเข้าชม Wongnai.com ในแต่ละวัน ปริมาณเพียง 2 ชั่วโมง")`
- `ax.set_xticks(x)`
- `ax.set_xticklabels(labels)`
- `ax.legend()`

19



20