Database Management
Homework 1

B11705044
Yen-Hung, Chiang

1

(a)

Weathersit	Days	Mean(temp)	Median(temp)	Std(temp)	Mean(cnt)
1	463	20.97	21.39	7.84	4876.79
2	247	19.29	18.79	6.85	4035.86
3	21	17.77	18.04	5.39	1803.29

Table 1: 各組別的統計量

```python
import pandas as pd
import openpyxl

df = pd.read_excel('Bike.xlsx', engine='openpyxl')
group1 = df.groupby('weathersit')

group1_size = group1.size()
group1_mean = group1['temp'].mean().round(2)
group1_median = group1['temp'].median().round(2)
group1_std = group1['temp'].std().round(2)
group1_mean_cnt = group1['cnt'].mean().round(2)

summary_df1 = pd.DataFrame({
    'Size': group1_size,
    'Mean Temperature': group1_mean,
    'Median Temperature': group1_median,
    'Standard Deviation Temperature': group1_std,
    'Mean Count': group1_mean_cnt
}).reset_index()

print(summary_df1)
```

Listing 1: (a) 小題的 Python Code

(b)

Weathersit	Workday	Days	Mean(temp)	Median(temp)	Std(temp)	Mean(cnt)
1	0	156	20.19	20.10	8.01	4587.27
	1	307	21.37	21.94	7.73	5023.90
2	0	70	18.99	17.72	6.93	3936.83
	1	177	19.40	19.03	6.84	4075.03
3	0	5	15.59	16.26	6.14	1815.40
	1	16	18.45	18.54	5.16	1799.50

Table 2: 各組別的統計量

```python
import pandas as pd
import openpyxl

df = pd.read_excel('Bike.xlsx', engine='openpyxl')
group2 = df.groupby(['weathersit', 'workday'])

group2_size = group2.size()
group2_mean = group2['temp'].mean().round(2)
group2_median = group2['temp'].median().round(2)
group2_std = group2['temp'].std().round(2)
group2_mean_cnt = group2['cnt'].mean().round(2)

summary_df2 = pd.DataFrame({
    'Size':group2_size,
    'Mean Temperature': group2_mean,
    'Median Temperature': group2_median,
    'Standard Deviation Temperature': group2_std,
    'Mean Count': group2_mean_cnt
}).reset_index()

print(summary_df2)
```

Listing 2: (b) 小題的 Python Code

(c)

Weathersit	corr. coef.
1	0.622190
2	0.644899
3	0.606836

Table 3: 三組資料各組內的 temp 和 cnt 的相關係數

```python
def compute_corr(group):
    if len(group) > 1:
        return group[['temp', 'cnt']].corr().iloc[0, 1]
    else:
        return float('nan')

corr_df = group1.apply(compute_corr).reset_index(name='Correlation')

print(corr_df)
```

Listing 3: (c) 小題的 Python Code

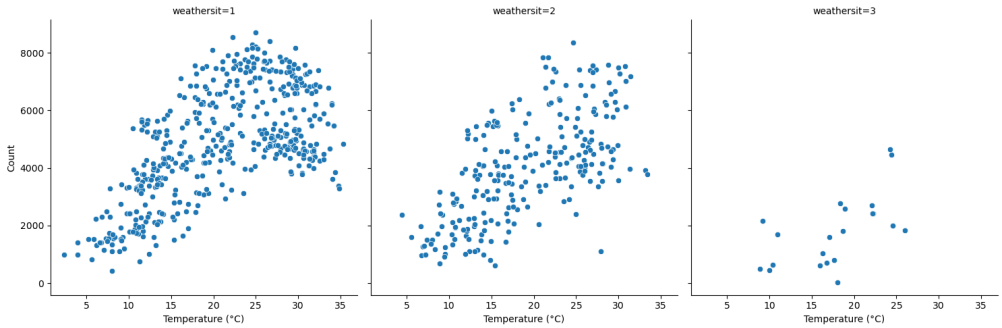


Figure 1: 各組資料 temp 和 cnt 的散佈圖。

```python
import seaborn as sns
import matplotlib.pyplot as plt

g = sns.FacetGrid(df, col='weathersit', col_wrap=3, height=5)
g.map_dataframe(sns.scatterplot, x='temp', y='cnt')

g.set_axis_labels('temp', 'cnt')
g.set_titles(col_template='weathersit={col_name}')
plt.show()
```

Listing 4: 產生散佈圖的 Python Code

2

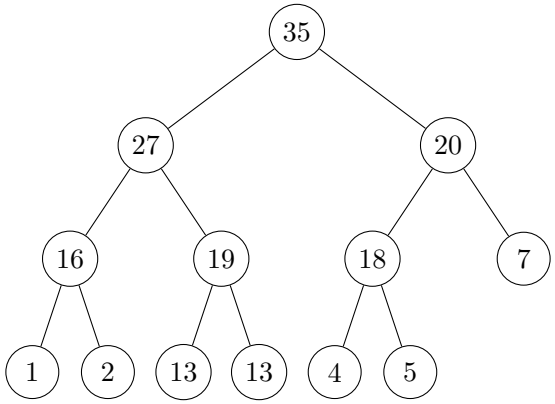
(a)

Assume there is a member function called `size()` in `class LinkedList`. It returns the size of the `LinkedList`.

- `constructor`: Set `LinkedList` to `nullptr`.
- `enqueue()`: Use `insert()` to add a node at the bottom of the list.
- `dequeue()`: Before dequeuing, check if the queue is empty.

  - If `isEmpty()` returns `True`, do nothing.
  - If `isEmpty()` returns `False`, use `delete()` to remove the node at the front of the list.
- `isEmpty()`: If `size()` is 0, then `isEmpty()` returns `True`. Otherwise, it returns `False`.
- `getQueueLength()`: Return `size()`.

(b)



(c)

Root Node:

Store the root node at index 0 in an array.

Leaf Nodes:

Suppose a node stored at index i in an array:

- The left child is stored at index $2i + 1$.
- The right child is stored at index $2i + 2$.

index	0	1	2	3	4	5	6	7	8	9	10	11	12
value	35	27	20	16	19	18	7	1	2	13	13	4	5

Table 4: Array of Max Heap

3

(a)

作業系統通常涵蓋 System Structures、Process Concept、Multithreaded Programming、Process Scheduling、Synchronization、Deadlocks、Memory-Management Strategies、Virtual-Memory Management、Thrashing、File System、I/O Systems。

(b)

Process 有獨立的 memory space 以及一個或以上的 thread，通常會把 process 劃分成多個 thread，每個 thread 是執行的基本單位，同個 process 裡面的 thread 共享 process 的 memory space。

(c)

連續分配原則是當有新的 Process 要儲存的時候，系統會在記憶體上找一塊夠大且連續的記憶體儲存。
優點：

- 存取快速
- 實現方式簡單

缺點：

- 會有 External Fragmentation 產生，讓記憶體的利用效率變低。
- 查找可用記憶體空間的時間成本高。

(d)

電腦在運行的時候常常會產生記憶體碎片化的情形，導致沒有一塊足夠大連續記憶體可以使用，為了解決這個問題所以產生了「虛擬記憶體」這項技術。這項技術主要的用途是在運行 Program 時不必將整個 Program 都載入到記憶體裡，只把常用的地方載入到記憶體就好。