

STAT 656 Final Exam

E. Lee Rainwater

4 August 2020

Mays Business School
Texas A&M University
College Station TX

Abstract

Short abstract here.

Figure 2: Corpus Sentiment Word Cloud

1 Summary

Overall Corpus Average Sentiment: 1.25

Refer to Table 1.

Table 1: Sentiment by Hotel

<i>hotel</i>	<i>doc</i>	<i>sentiment</i>	<i>n_words</i>
Bally's	162.5	1.205211	13.268519
Bellagio	493.0	1.709758	14.270030
Circus Circus	832.5	0.907237	12.567251
Encore	1170.5	1.353040	17.565868
Excalibur	1504.5	1.088394	14.302395

2 Code Listing

Below is the SAS code used to generate the tabulation of wine points and cost.

```
1 ODS RTF FILE="C:\Users\rainwater-e\...\hw-08\CABERNET.RTF";
2 PROC TABULATE DATA=&EM_IMPORT_DATA;
3 CLASS TextCluster_cluster_;
4 VAR POINTS PRICE;
5 TABLE textCluster_cluster_, (POINTS PRICE)*MEAN;
6
7 RUN;
8 ODS RTF CLOSE;
```

3 Word Clouds

Figure 1: Corpus Total Word Cloud

Figure 3: Bellagio Total Word Cloud

Figure 4: Bellagio Total Word Cloud

4 Code Output

```
Python 3.7.6 | packaged by conda-forge | (default, Jun 1 2020, 18:11:50) [MSC
v.1916 64 bit (AMD64)]
```

```
Type "copyright", "credits" or "license" for more information.
```

```
IPython 7.16.1 -- An enhanced Interactive Python.
```

```
In [1]: runfile('C:/Users/rainwater-e/OneDrive - Texas A&M University/Summer-2020/stat656/hw-10')
wdir='C:/Users/rainwater-e/OneDrive - Texas A&M University/Summer-2020/stat656/hw-10')
```

```
*****
***** READING DATA SOURCE... *****
*****
```

```
TEXT LENGTH:
Max. Accepted 32000
Max. Observed 11075
Truncated 0
```

```
*****
***** CREATING TOKEN COUNT MATRIX... *****
*****
```

```
*****
***** CALCULATE AND STORE SENTIMENT SCORES... *****
*****
```

Corpus Average Sentiment: 1.2521789531013003

Most Negative with 2 or more Sentiment Words:

Review 57 Sentiment is -2.75

Review 160 Sentiment is -2.75

Most Positive with 2 or more Sentiment Words:

Review 692 Sentiment is 4.50

Maximum Text Length---- 11075

Total Reviews----- 1671

Total Sentiment Terms-- 24050

Avg. Sentiment Terms--- 14.39

hotel sentiment n_words

0 Bally's 2.000000 7

1 Bally's 1.533333 15

2 Bally's 1.928571 14

3 Bally's 1.125000 24

4 Bally's -1.222222 9

... ..

1666 Excalibur 2.416667 12

1667 Excalibur 1.705882 17

1668 Excalibur 1.600000 5

1669 Excalibur 1.100000 20

1670 Excalibur 0.625000 8

[1671 rows x 3 columns]

doc sentiment n_words

hotel

Bally's 162.5 1.205211 13.268519

Bellagio 493.0 1.709758 14.270030

Circus Circus 832.5 0.907237 12.567251

Encore 1170.5 1.353040 17.565868
Excalibur 1504.5 1.088394 14.302395

```
*****
***** GENERATING TOTAL WORD CLOUD FOR CORPUS... *****
*****
```

Figures now render in the Plots pane by default. To make them also appear inline in the Console, uncheck "Mute Inline Plotting" under the Plots pane options menu.

<Figure size 432x288 with 0 Axes>

```
*****
***** GENERATING SENTIMENT WORD CLOUD FOR CORPUS... *****
*****
<Figure size 432x288 with 0 Axes>
```

SENTIMENT TERMS

Unique.... 1361
Total..... 24050

```
*****
***** GENERATING TOTAL WORD CLOUD FOR BELLAGIO... *****
*****
<Figure size 432x288 with 0 Axes>
```

```
*****
***** GENERATING SENTIMENT WORD CLOUD FOR BELLAGIO... *****
```

***** CALCULATE AND STORE SENTIMENT SCORES FOR BELLAGIO... *****

Corpus Average Sentiment: 1.7097579915456067

Most Negative with 2 or more Sentiment Words:
Review 32 Sentiment is -1.00

Most Positive with 2 or more Sentiment Words:
Review 298 Sentiment is 4.33
<Figure size 432x288 with 0 Axes>

BELLAGIO SENTIMENT TERMS

Unique.... 647
Total..... 4809
