STA 440 Individual Case Study

Mitchelle Mojekwu

2023-12-14

Introduction

These are the three main objectives for this study:

1) Is there an association between 30-day mortality and the day of the week in which an operation occurs?
2) Is there an association between 30-day mortality and the time of day in which an operation occurs?
3) Is there an association between 30-day mortality and the month of the year in which an operation occurs?

Exploratory Data Analysis

Summary Statistics of Baseline Risk Factors for Patients

Predictors	Summary.Statistics
Age(mean/SD)	
Age	57.66 ± 15.04
Gender (count/%)	
Male	14768 (46.15 %)
Female	17230 (53.84 %)
Race (count/%)	
Caucasian	26488 (82.77 %)
African American	3790 (11.84 %)
Other	1243 (3.88 %)
ASA Physical Status (count/%)	
I-II	17261 (53.94 %)
III	13677 (42.74 %)
IV-VI	1055 (3.3 %)
BMI (median/IQR)	
Body Mass Index	28.19 [24.6 , 32.81]
Diseases/Disorders (count/%)	
Cancer	10958 (34.24 %)
Cardiovascular/cerebrovascular disease	16176 (50.55 %)
Dementia	242 (0.76 %)
Diabetes	4166 (13.02 %)
Digestive disease	7037 (21.99 %)
Osteoarthritis	5719 (17.87 %)
Psychiatric disorder	2910 (9.09 %)
Pulmonary disease	3493 (10.92 %)
Indexes (median/IQR)	
Charlson Comorbidity Index	0 [0 , 2]
Risk Stratification Index (30-day mortality)	-0.3 [-1.24 , 0]
Risk Stratification Index (in-hospital complications)	-0.27 [-0.84 , 0]

Summary Statistics of Case Frequencies (Count/%)

Predictors	Summary.Statistics
Hour of Operation	
6:00	562 (1.76 %)
7:00	10631 (33.22 %)
8:00	3807 (11.9 %)
9:00	1664 (5.2 %)
10:00	2501 (7.82 %)
11:00	2855 (8.92 %)
12:00	2763 (8.63 %)
13:00	2623 (8.2 %)
14:00	2063 (6.45 %)
15:00	1267 (3.96 %)
16:00	745 (2.33 %)
17:00	356 (1.11 %)
18:00	163 (0.51 %)
19:00	1 (0 %)
Day of the Week	
Monday	7005 (21.89 %)
Tuesday	7008 (21.9 %)
Wednesday	6266 (19.58 %)
Thursday	5635 (17.61 %)
Friday	6087 (19.02 %)
Month	
January	2670 (8.34 %)
February	2506 (7.83 %)
March	2697 (8.43 %)
April	2698 (8.43 %)
May	2654 (8.29 %)
June	2994 (9.36 %)
July	2325 (7.27 %)
August	3177 (9.93 %)
September	3208 (10.02 %)
October	2689 (8.4 %)
November	2544 (7.95 %)
December	1839 (5.75 %)
Phase of Moon	
New Moon	7708 (24.09 %)
First Quarter	8100 (25.31 %)
Full Moon	8051 (25.16 %)
Last Quarter	8142 (25.44 %)

Methodology

```
model <- glm(mort30 ~ age+hour+dow+month+ moonphase, data = SurgeryTiming, family = "binomial")

# Display the summary of the model
summary(model)
```

```
##
## Call:
## glm(formula = mort30 ~ age + hour + dow + month + moonphase,
##     family = "binomial", data = SurgeryTiming)
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -10.187427   0.661003 -15.412  < 2e-16 ***
## age                    0.054765   0.006756   8.106 5.21e-16 ***
## hour                   0.129940   0.027345   4.752 2.02e-06 ***
## dowMonday             -0.182428   0.257806  -0.708   0.4792
## dowThursday           -0.183813   0.273734  -0.672   0.5019
## dowTuesday            -0.100698   0.253527  -0.397   0.6912
## dowWednesday          -0.535081   0.284700  -1.879   0.0602 .
## monthAugust           -0.478068   0.442619  -1.080   0.2801
```

```
## monthDecember          -0.664208   0.579286  -1.147   0.2515
## monthFebruary           0.449082   0.379229   1.184   0.2363
## monthJanuary            0.482337   0.371067   1.300   0.1936
## monthJuly               0.110020   0.410410   0.268   0.7886
## monthJune               0.033914   0.395194   0.086   0.9316
## monthMarch              0.003232   0.410111   0.008   0.9937
## monthMay               -0.194284   0.430041  -0.452   0.6514
## monthNovember          -0.812205   0.533872  -1.521   0.1282
## monthOctober           -0.482478   0.458442  -1.052   0.2926
## monthSeptember          0.038993   0.384029   0.102   0.9191
## moonphaseFull Moon      0.200588   0.241455   0.831   0.4061
## moonphaseLast Quarter   0.154164   0.242774   0.635   0.5254
## moonphaseNew Moon       0.022567   0.253383   0.089   0.9290
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1778.6  on 31998  degrees of freedom
## Residual deviance: 1659.9  on 31978  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: 1701.9
##
## Number of Fisher Scoring iterations: 9
```

Results

Discussion

Appendix

Extra EDA Visuals

```
## Warning: Removed 2 rows containing non-finite values (`stat_density()`).
```

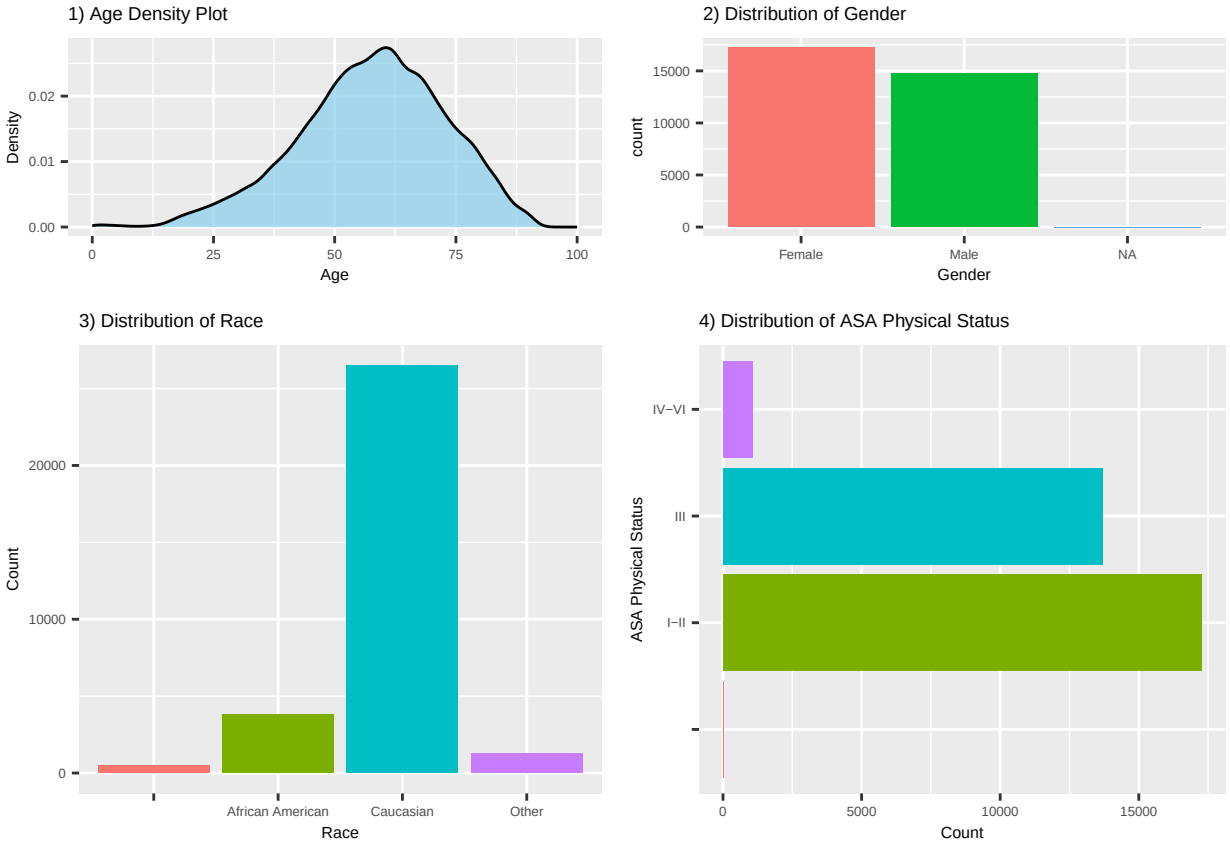


Image 1) Age

Since the distribution of the patients' age is relatively normal and unimodal and not really that skewed to the left the best summary statistic is the mean and the best metric of spread is the standard deviation.

Image 2) Gender

There appears to be a bit more operation cases done on women than men.

Image 3) Race

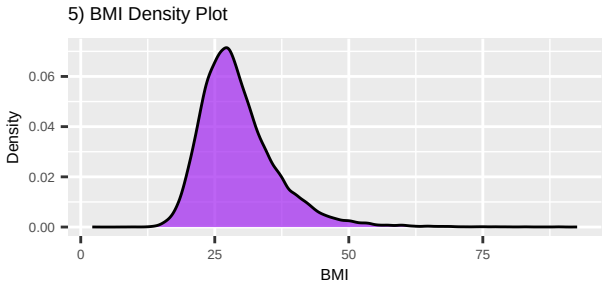
There are multidude more operation cases done on Caucasian patients than any other racial demographics.

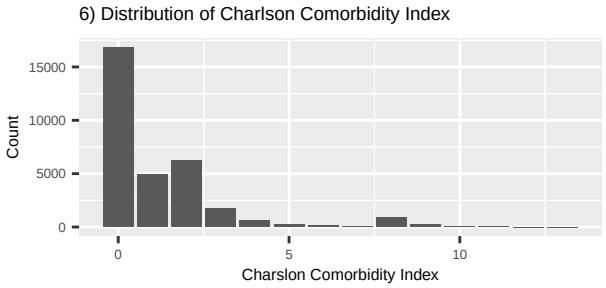
Image 4) ASA Physical Status

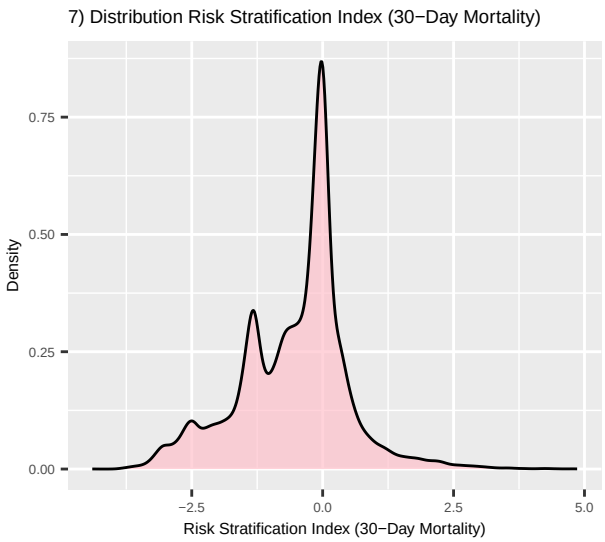
*Note: ASA classification uses a grading system of I (one) through V (five) I identifying a person in good health and V as a person with a severe, life-threatening condition. The sixth (VI) status identifies deceased organ donors.

Most patients of operation cases reside in the I-II ASA status and then the III status comes in a close second. There are not many patients that reside in the IV-VI status.

```
## Warning: Removed 3290 rows containing non-finite values (`stat_density()`).
```







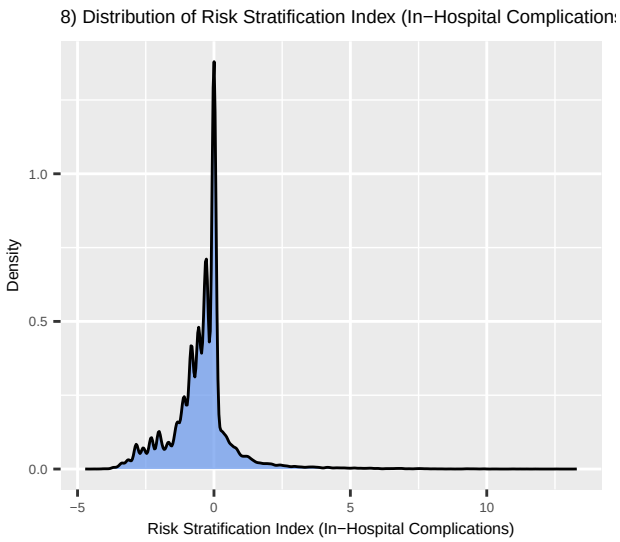


Image 5) BMI

Since the distribution of the patients' BMI is slightly skewed to the right the best summary statistic is the median and the best metric of spread is the IQR.

Image 6) Charlson Comorbidity Index

*Note: The Charlson comorbidity index (CCI) is the most widely used comorbidity index used to determine survival rate (1yr and 10yr) in patient with multiple comorbidities. As the CCI level increases the probability of one year of survival decreases.

Since the distribution of the patients' Charlson Comorbidity Index is skewed to the right the best summary statistic is the median and the best metric of spread is the IQR.

Image 7) Risk Stratification Index (30-Day Mortality)

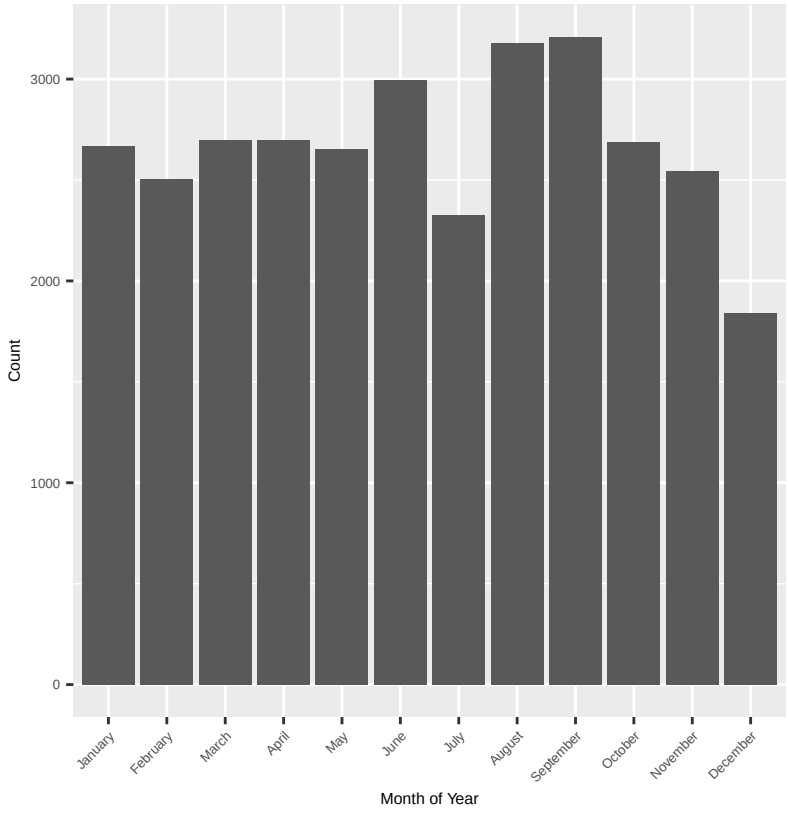
*Note: The Risk Stratification Index (RSI) is an open source, nationally validated, risk stratification methodology that permits outcomes such as duration of hospitalization and mortality to be compared equally across institutions.

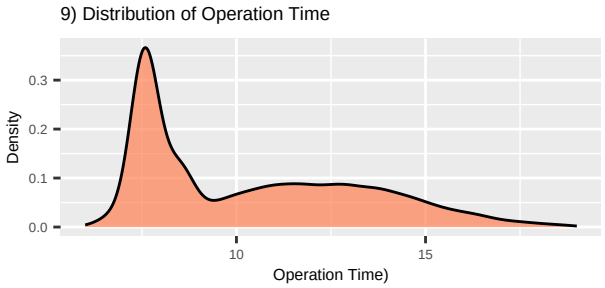
The density chart of the distribution of RSI (30-day mortality) is multimodal and skewed to the left so the median and IQR would be more accurate measures of center and spread.

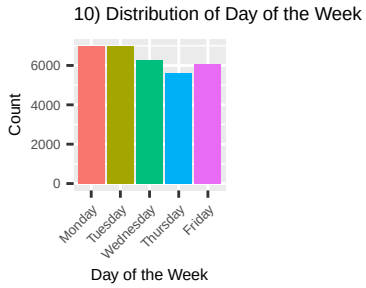
Image 8) Risk Stratification Index (In-Hospital Complications)

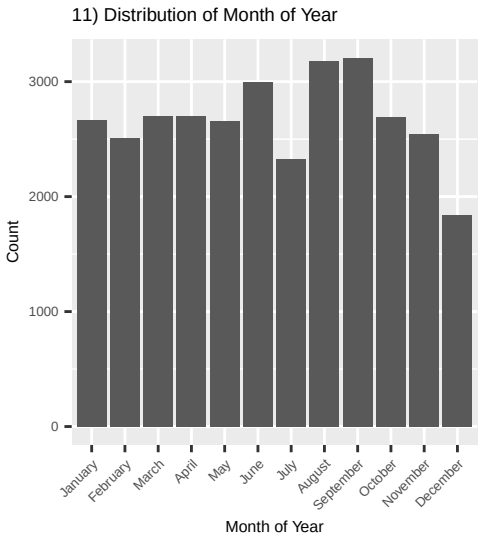
The density chart of the distribution of RSI (in-hospital complications) is multimodal and skewed to the left so the median and IQR would be more accurate measures of center and spread.

11) Distribution of Month of Year









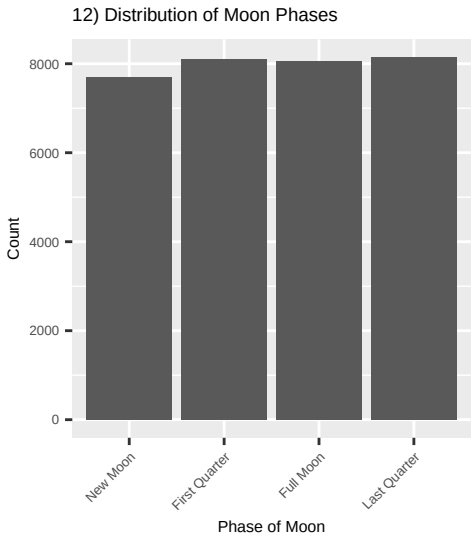


Image 9) Operation Hour

The distribution is bimodal with a huge concentration of operation hours at around 8:00 am and a much smaller concentration at 12:30 pm.

Image 10) Day of the Week

The distribution of the days the of operation cases tend have a higher concentration in the beginning of the week.

Image 11) Month of the Year

The distribution of the months of operation cases show peaks in June and near the end of the summer (August and September).

Image 12) Phase of Moon

The distribution of the moon phases of operation cases show a relatively uniform distribution with the least cases occuring during new moon phases.

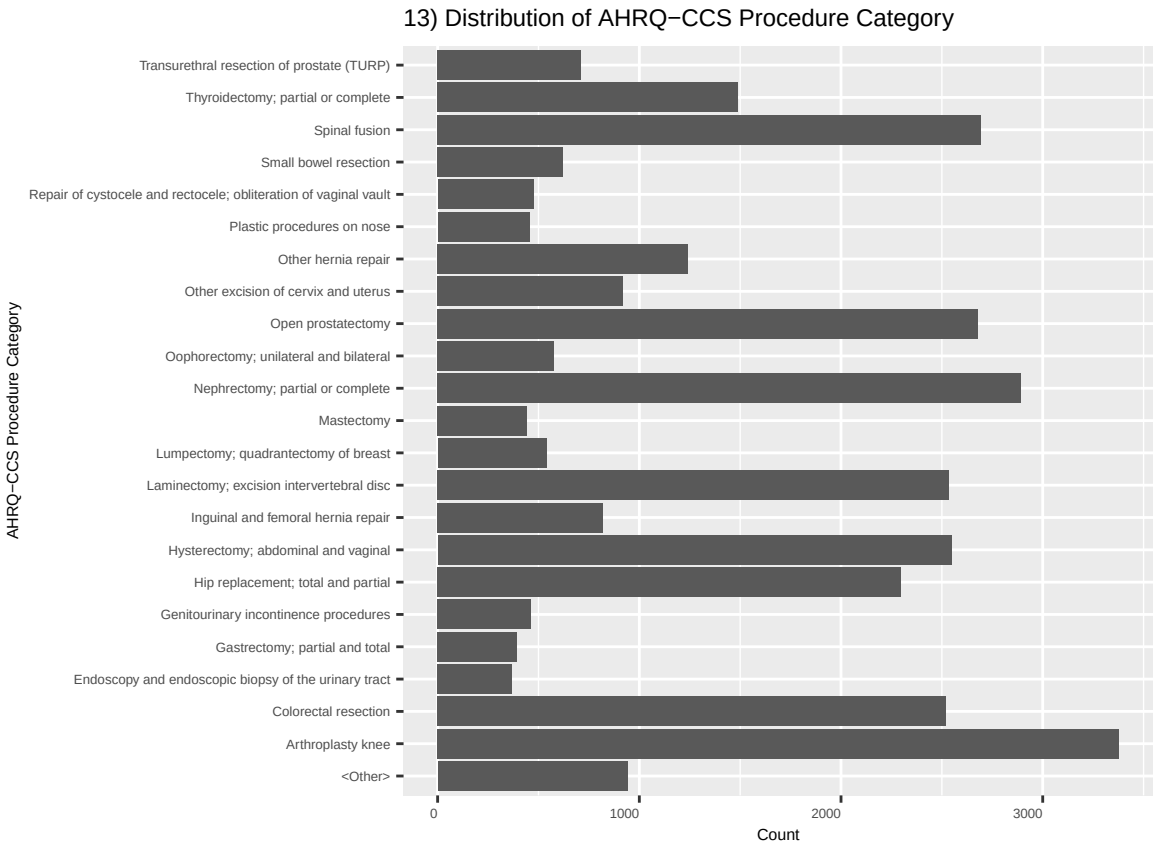


Image 13) AHRQ-CCS Procedure Category

The distribution of procedure type shows there are multiple sporadic procedure types that have very high concentrations of operations and multiple that have low concentrations.

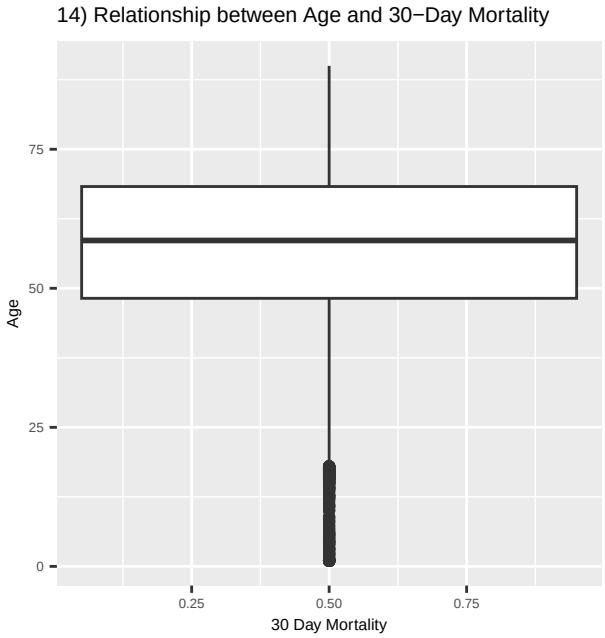
```
## Warning: Continuous x aesthetic
## i did you forget `aes(group = ...)`?
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_boxplot()`).
```

```
## Warning: The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
```

```
## Warning: Continuous x aesthetic
## i did you forget `aes(group = ...)`?
```

```
## Warning: The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
```



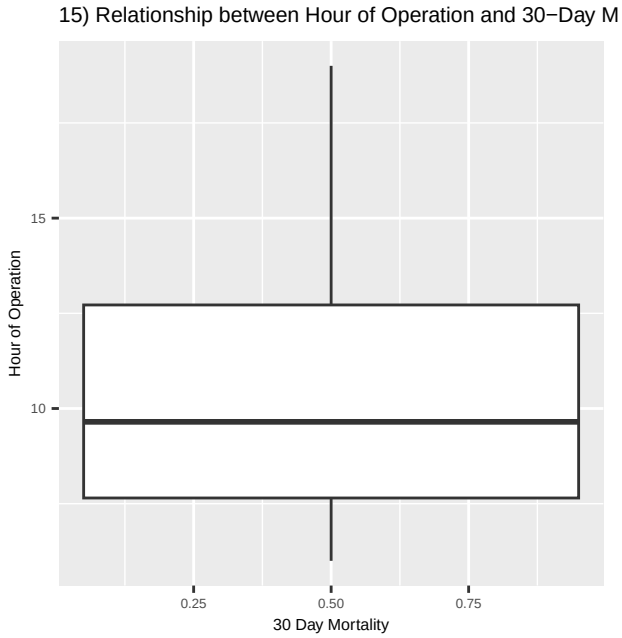


Image 14) Age and 30 Day Morality

The box plots show that operations of patients that have 30-day mortality are more generally of an older age than patients who do not have 30-day mortality.

Image 15) Hour and 30 Day Morality

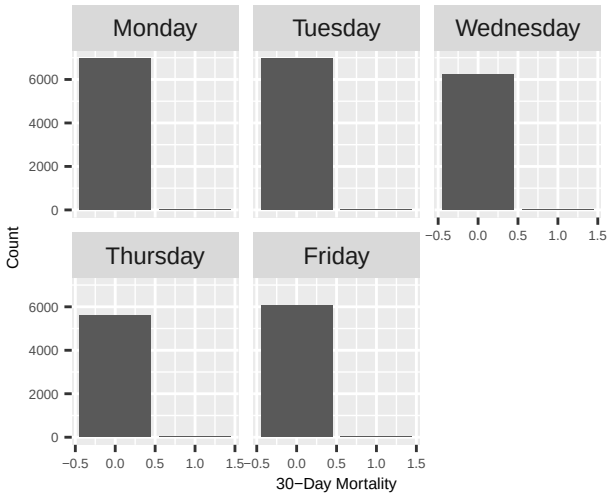
The box plots show that operations of patients that have 30-day mortality tend to happen later in the day than operations of patients that do not have 30-day mortality.

```
## Warning: The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
```

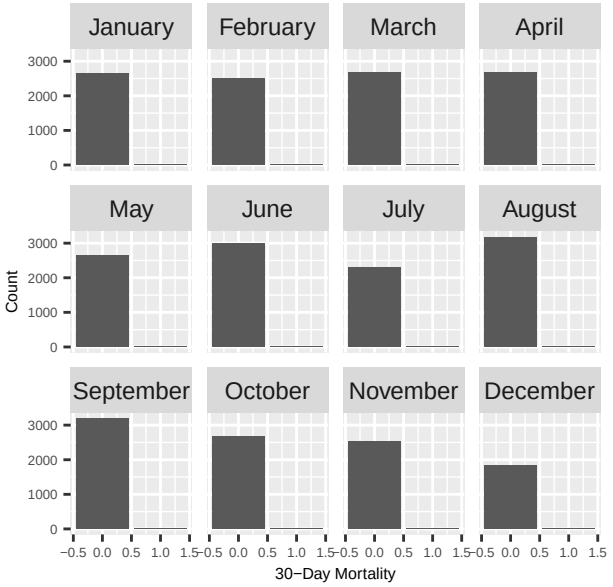
```
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
```

```
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: fill
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
```

16) Relationship between Day of Operation & 30−Day Mortality



17) Relationship between Month of Operation & 30−Day Mortality




```
## `summarise()` has grouped output by 'dow'. You can override using the `.groups`
## argument.
```

dow	mort30	Count
Monday	0	6975
Monday	1	30
Tuesday	0	6976
Tuesday	1	32
Wednesday	0	6245
Wednesday	1	21
Thursday	0	5611
Thursday	1	24
Friday	0	6056
Friday	1	31

```
## [1] 0.004282655
```

```
## [1] 0.00456621
```

```
## [1] 0.00335142
```

```
## [1] 0.004259095
```

```
## [1] 0.005092821
```

```
## `summarise()` has grouped output by 'month'. You can override using the
## `.groups` argument.
```

month	mort30	Count
January	0	2651
January	1	19
February	0	2489
February	1	17
March	0	2685
March	1	12
April	0	2686
April	1	12
May	0	2644
May	1	10
June	0	2980
June	1	14
July	0	2313
July	1	12
August	0	3168
August	1	9
September	0	3192
September	1	16
October	0	2681
October	1	8
November	0	2539
November	1	5
December	0	1835
December	1	4

```
## [1] 0.007116105
```

```
## [1] 0.006783719
```

```
## [1] 0.004449388
```

```
## [1] 0.004447739
```

```
## [1] 0.003767898
## [1] 0.004676019
## [1] 0.00516129
## [1] 0.002832861
## [1] 0.004987531
## [1] 0.002975084
## [1] 0.001965409
## [1] 0.002175095
```

Image 16) Day of Week and 30 Day Mortality

The day with the highest proportion of 30-day mortality cases is Friday, then Tuesday, then Monday, then Thursday, then Wednesday.

Image 17) Month and 30 Day Mortality

The month with the highest proportion of 30-day mortality cases is January, then February, then July.

Work Cited