

# HOW MUCH TROUBLE IS A LITTLE OVERCROWDING?

An analysis by Quinton Carter



# HOSPITAL OVERCROWDING

- A Hospitals ED is measured by a system called NEDOC



- ED Overcrowding dangers
  - patient satisfaction scores
  - patient care
  - patient's symptoms
  - potentially a clinical outcome
- Many theories & hypotheses, have proven effective against Overcrowding
- One definitive solution to the issue has yet to be found.



# GOAL

- Identification of ED trends and creation of predictive analytics to combat future overcrowding
- Data Visual interpretations through from R and Tableau IDE
- Predictive models created and benefit
  - Support Vector Machine
  - Naïve Bayes
  - Linear Regression
  - Random Forest
- This analysis will center around overcrowding
  - NEDOC average greater than 101
  - NEDOC score of 2 and above.

Above 180  
Disaster

4

141-180  
Severe

3

101-140  
Overcrowded

2

51-100  
Busy

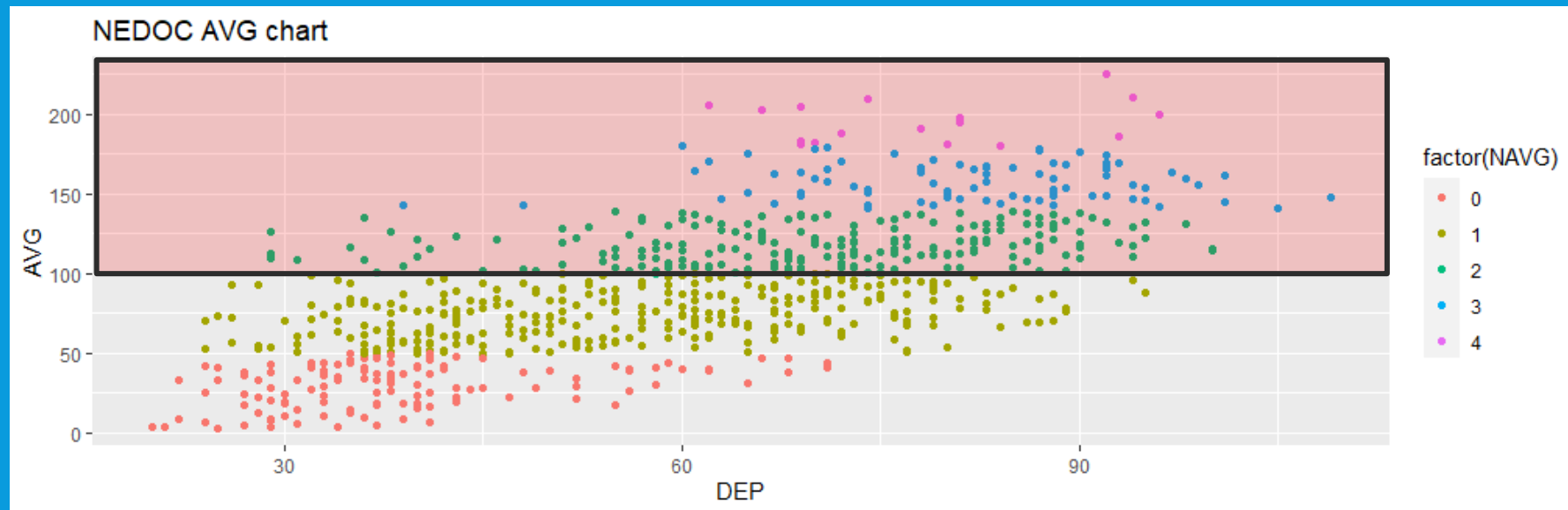
1

0-50  
Normal

0

# DISTRIBUTION OF NEDOC SCORES

- DEP increases in correlation to NEDOC Score increase
- NEDOC Scores from 0 to 3 are evenly distributed.
- Score of 4 occurs but is minimal.
- The area of focus is ED overcrowding, which occurs when NEDOC Score is 2 or more
  - 2 – OVR (overcrowded)
  - 3 – SEV (severe)
  - 4 – DNG (disastrous)

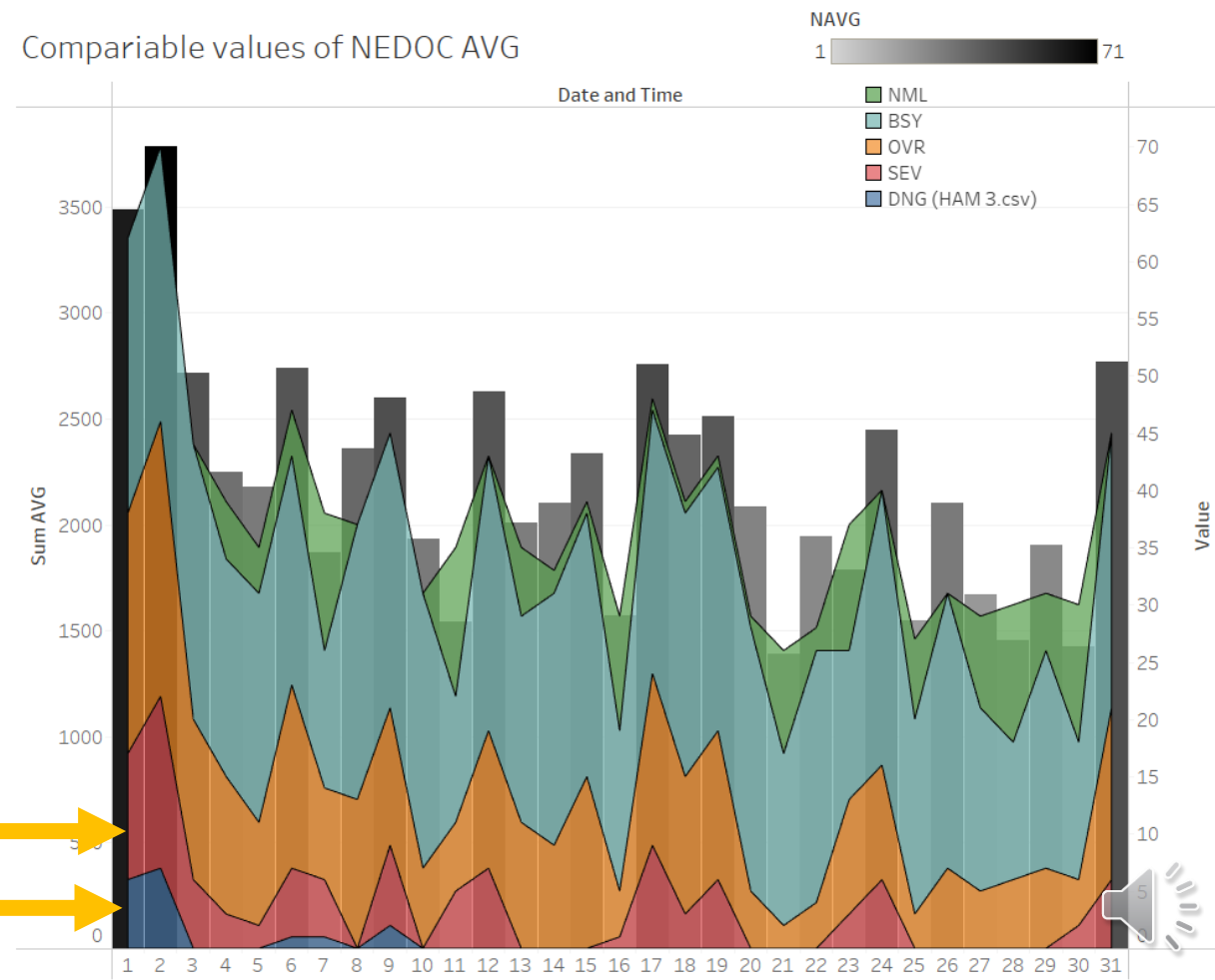


# DISTRIBUTION - ADDITIONAL VIEWS

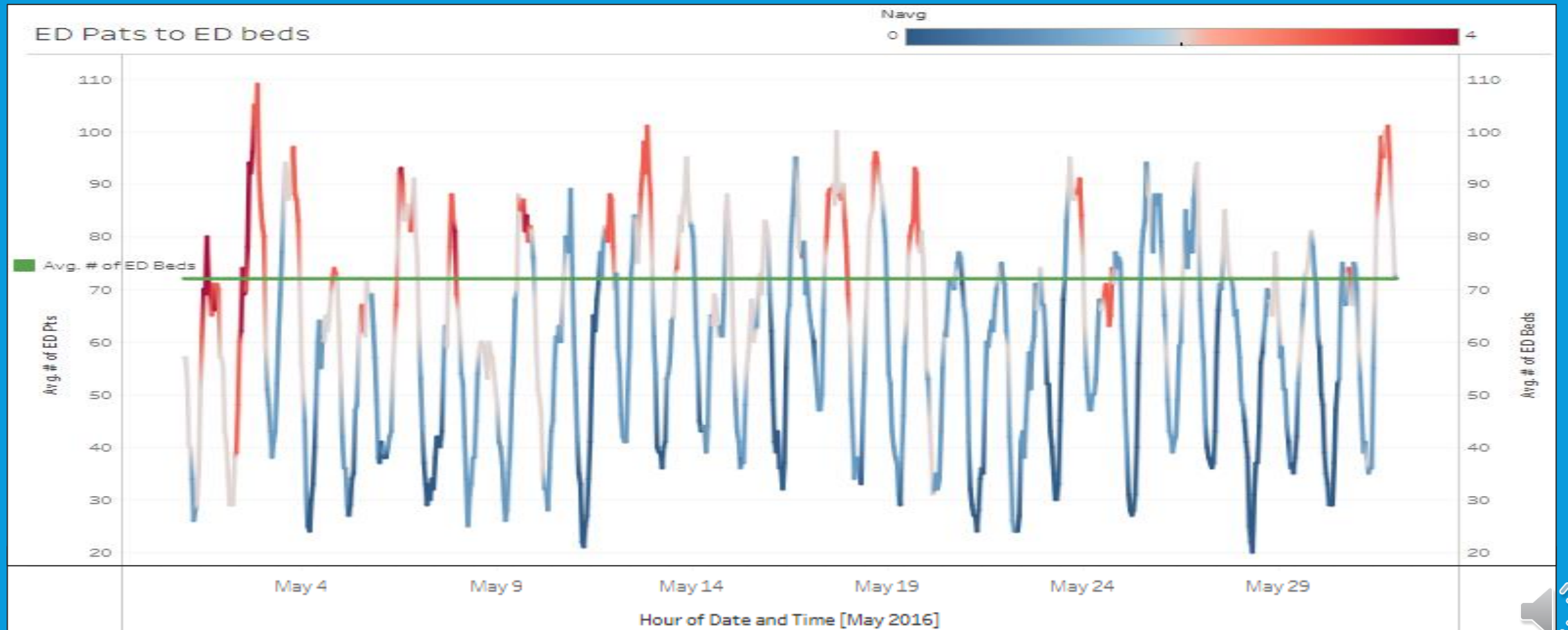
Summation of NEDOC scores



Comparable values of NEDOC AVG

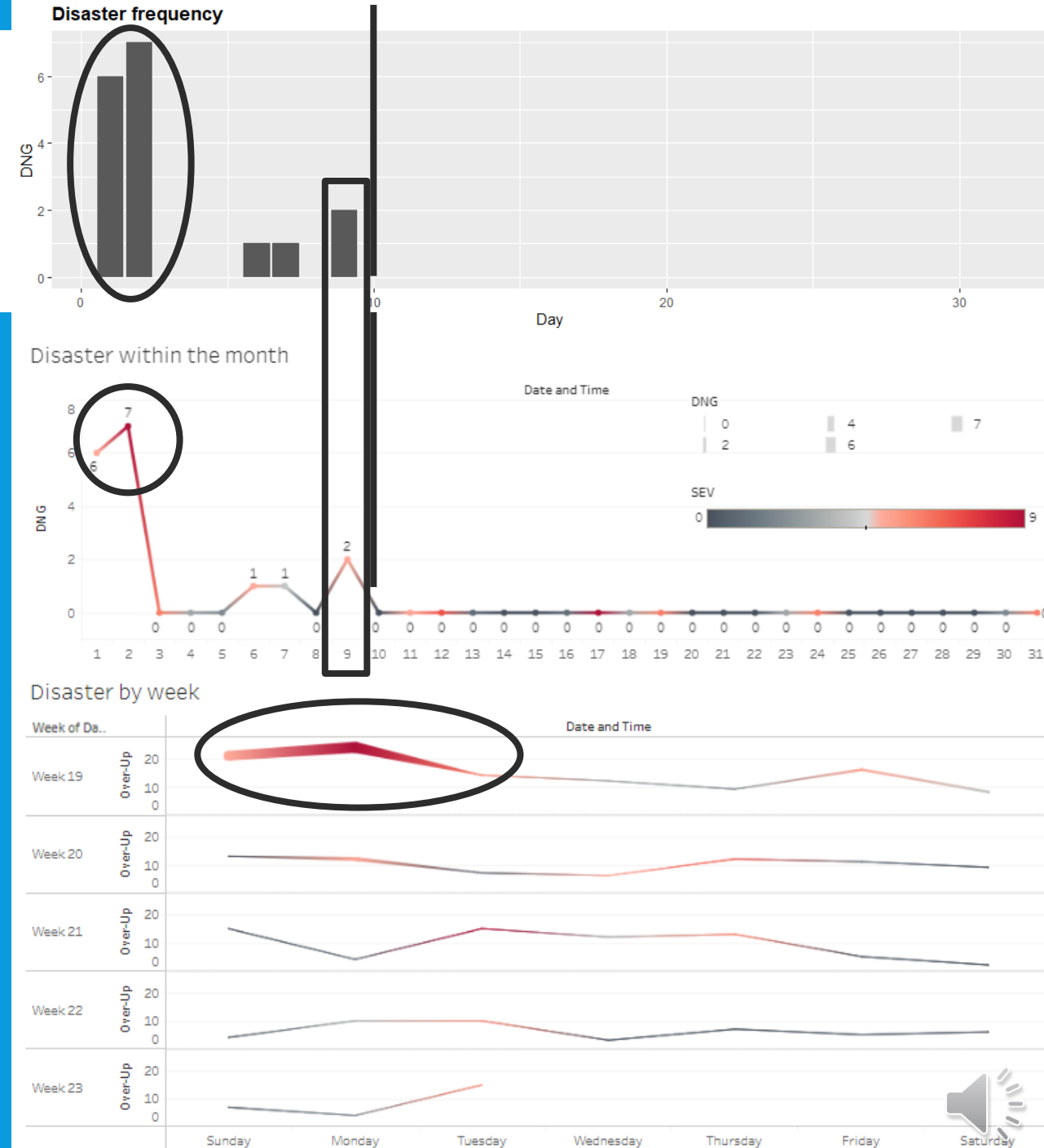


# CONSTANT VARIABLE



# DISASTER ANALYSIS

- Overcrowding occurred 311 times out of 744 data points.
- Overcrowded and Severe variables were well distributed
  - Overcrowded – 206
  - Severe – 88
- The Disaster score is not well distributed
  - Occurred less than 17
  - 13 occurrences in the first 2 days
- For these reasons, the Disaster score can be interpreted as an outlier.

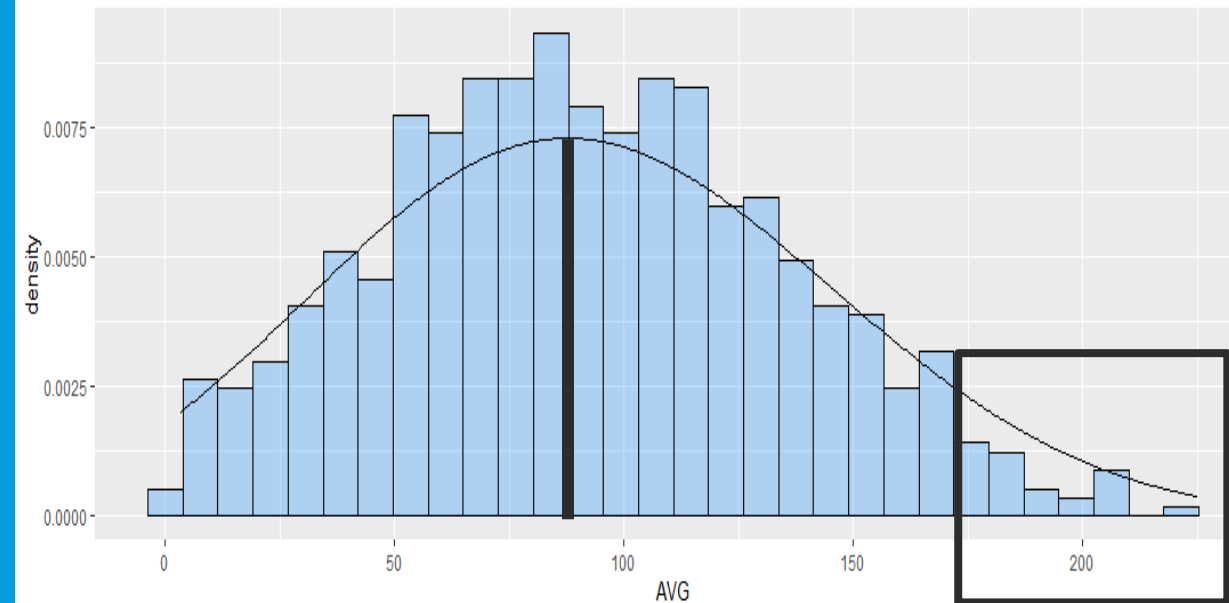


# DISASTER ANALYSIS 2

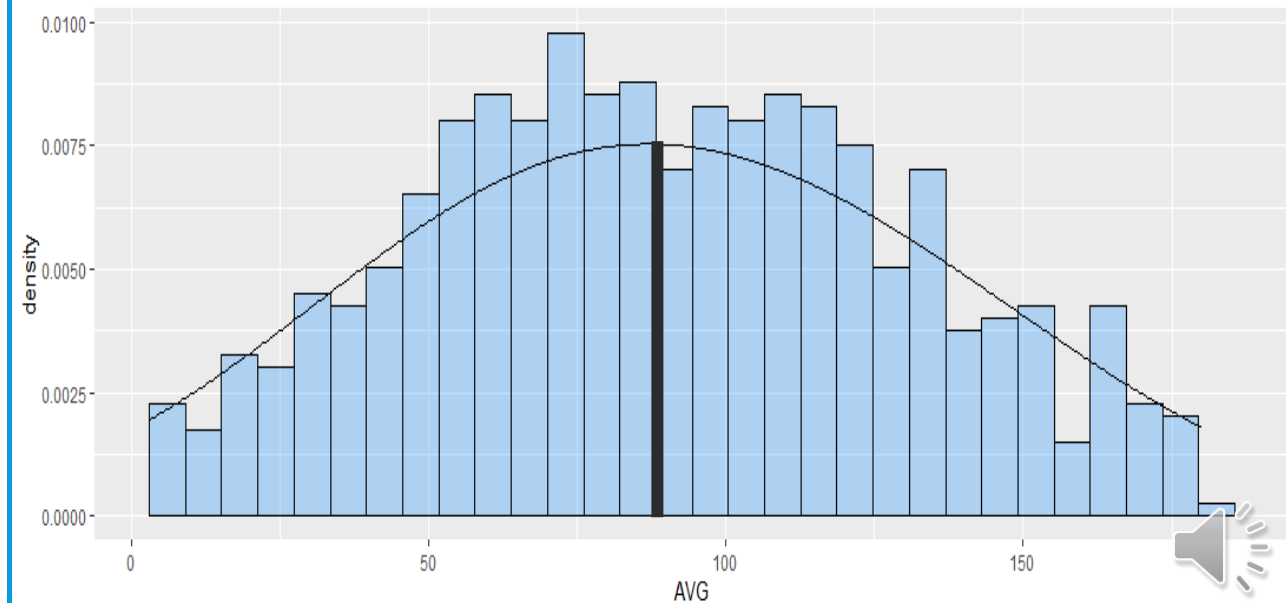
Below are distribution charts of NEDOC averages that relate to the disaster variable

- If disaster remains within the data set the top end of the distribution is highly unbalanced.
- In contrast when disaster is excluded, the scores are much more balanced and evenly distributed.

Distribution of the NeDoc Averages



NE\_Week Distribution of the NeDoc Averages

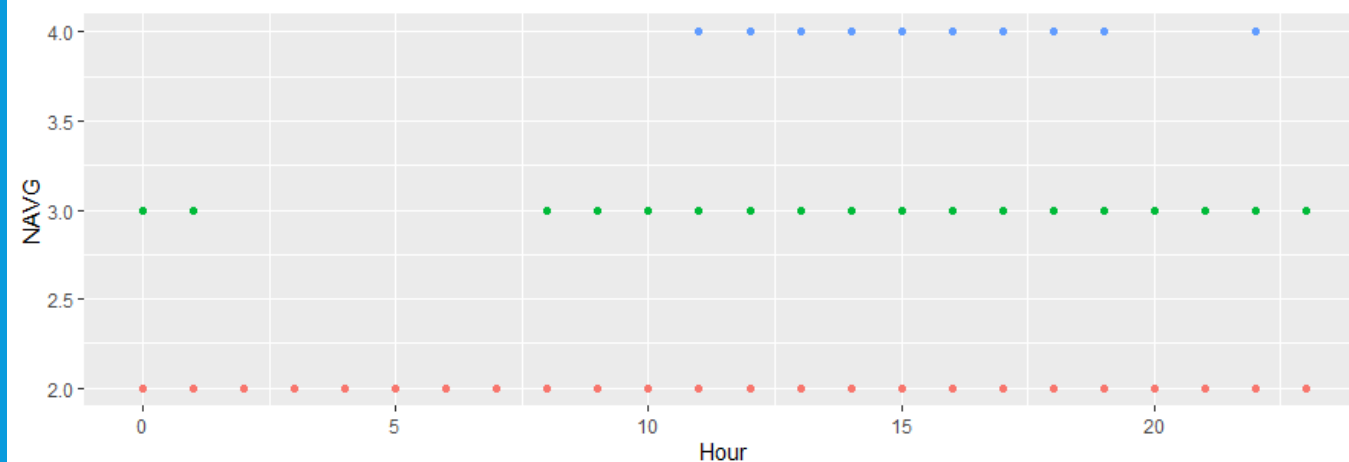




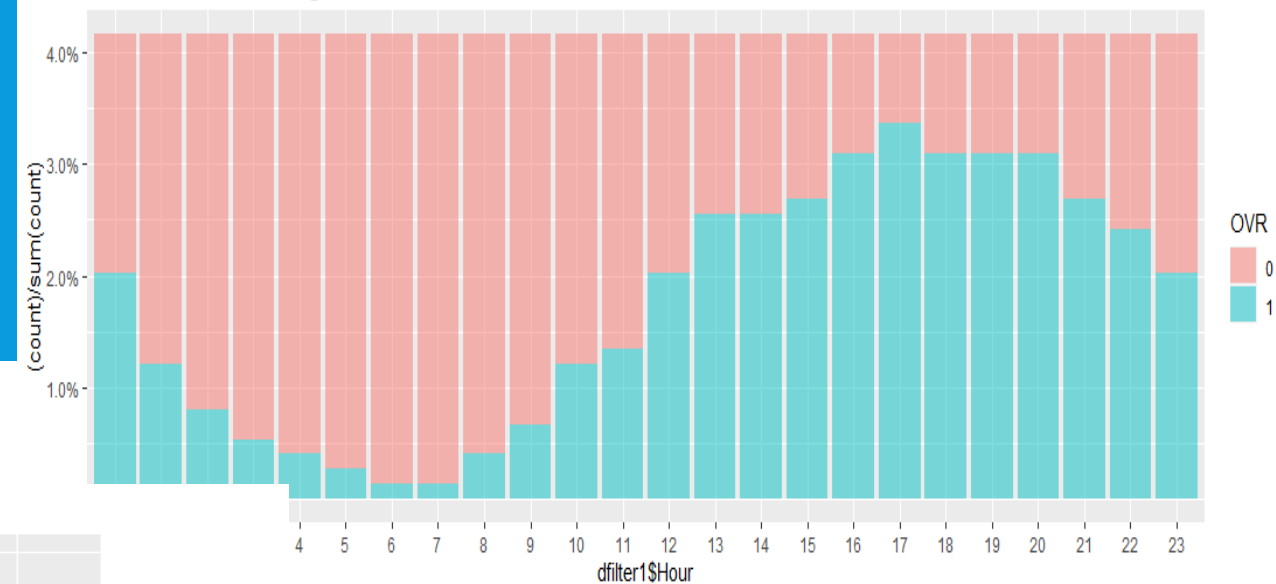
# NEDOC SCORE HOUR DISTRIBUTION

- Hour is a good indicator of higher NEDOC scores.
- ED Hours between 10 am and 0 (midnight) have higher NEDOC scores.
- Between these hours NEDOC values from 1 to 3 occur
  - Normal score or 0/Orange appears as an outlier

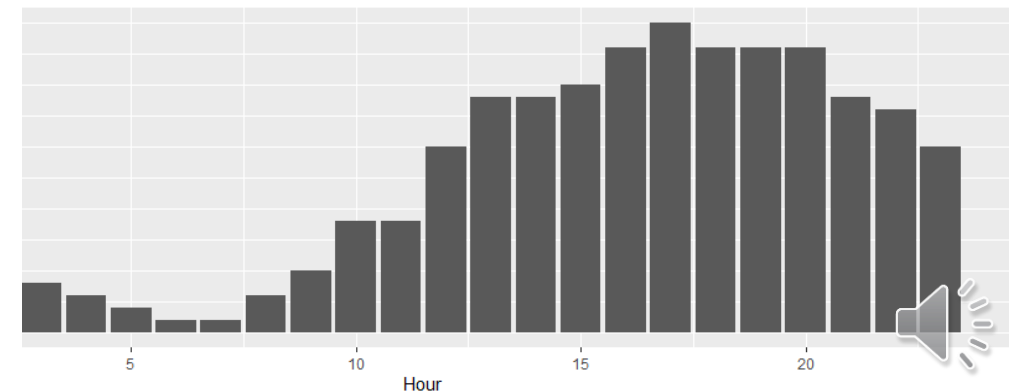
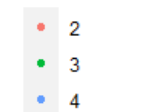
NEDOC Disbursement



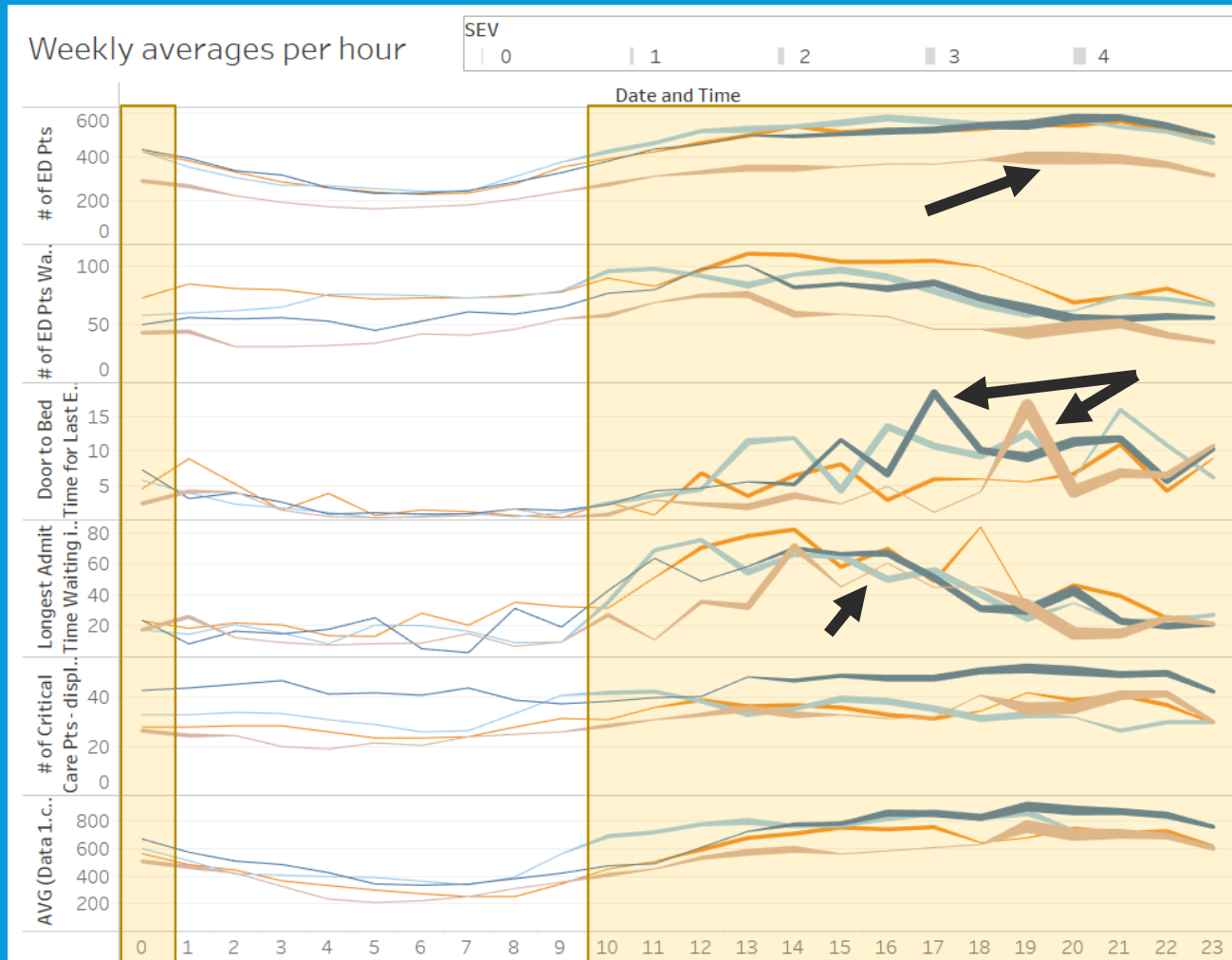
Hour for Overcrowding



factor(NAVG)



# NEDOC VARIABLE HOUR DISTRIBUTION P<sub>2</sub>

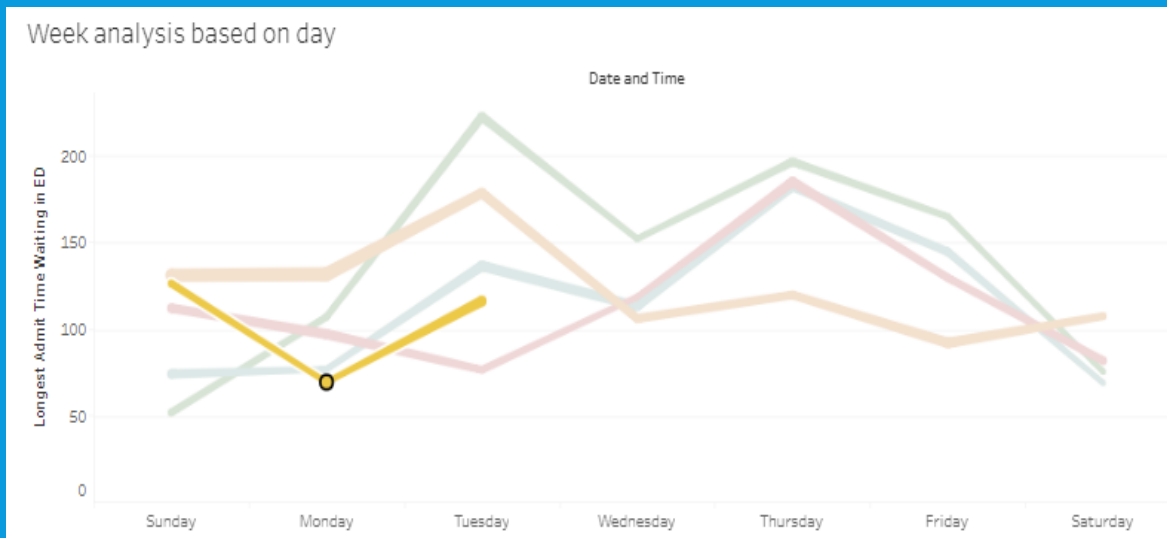


- This chart compares each primary variable to Hour and includes a line for each week of data
- The peak hours have a dramatic effect on each variables, week over week routine.
- Severity (line thickness) predominately appears within the peak hours
- ED peak hours cause all variables to experience volatile explosions or sporadic behavior



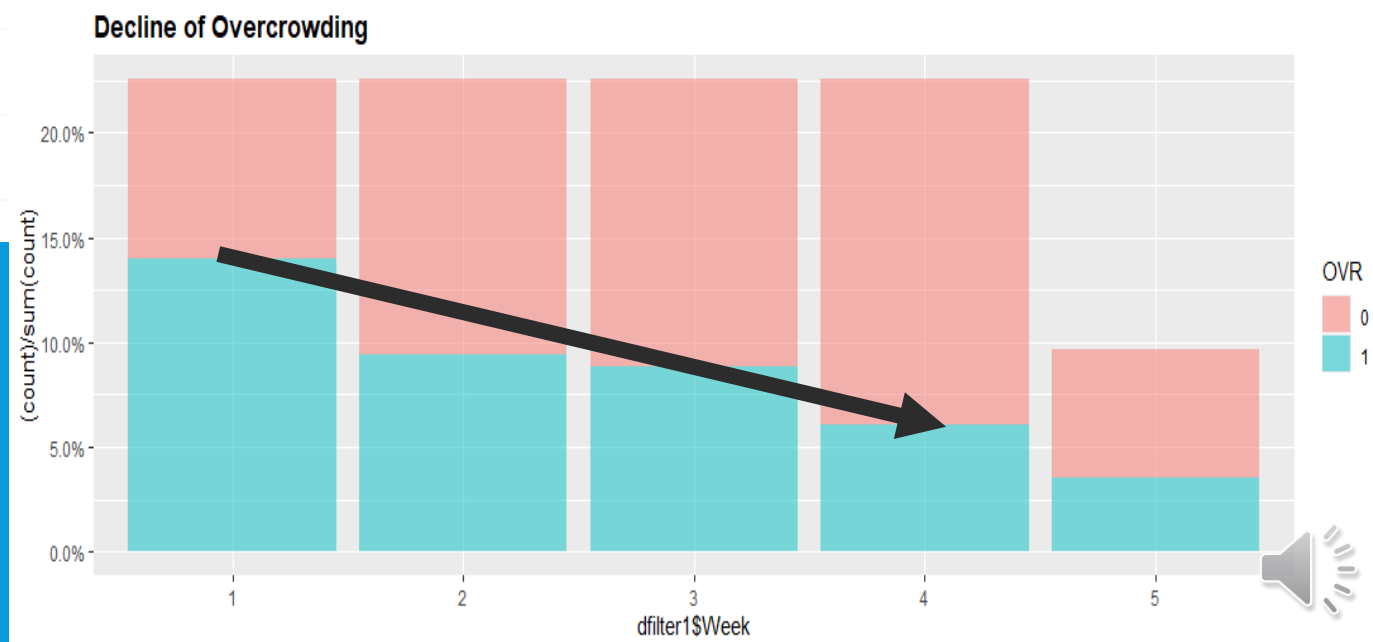
# WEEK ANALYSIS

These weekly charts show macro views of the data and trends.



- Something not standard occurs in week 5.
- Identified that this is not a full week of data, which could skew results.

- This chart indicates the ED's progression to manageable NEDOC score.
- Could the ED have already implemented something to help them manage their NEDOC score?

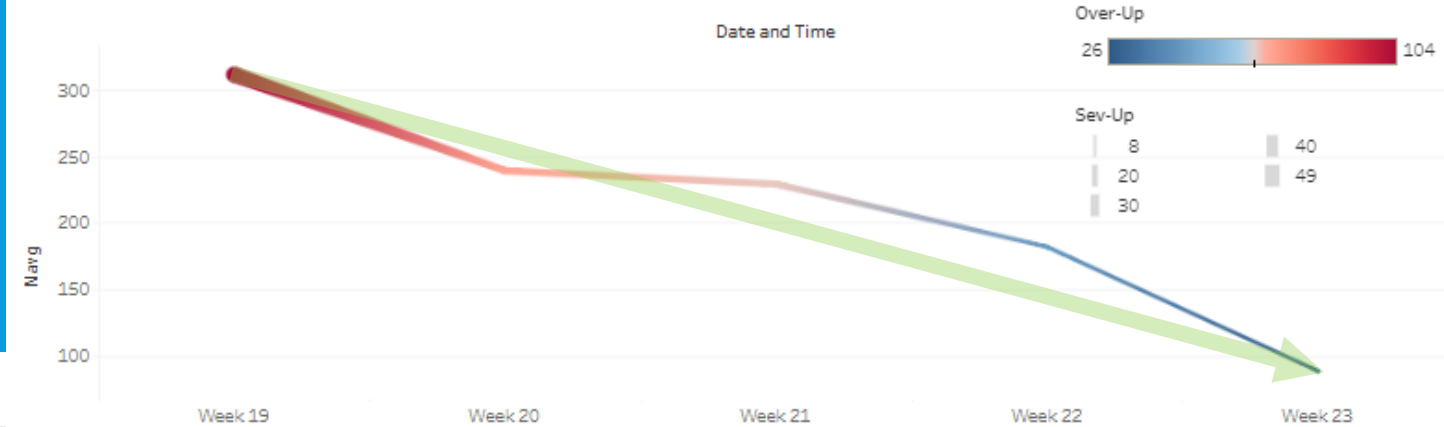


# WEEK ANALYSIS P<sub>3</sub>

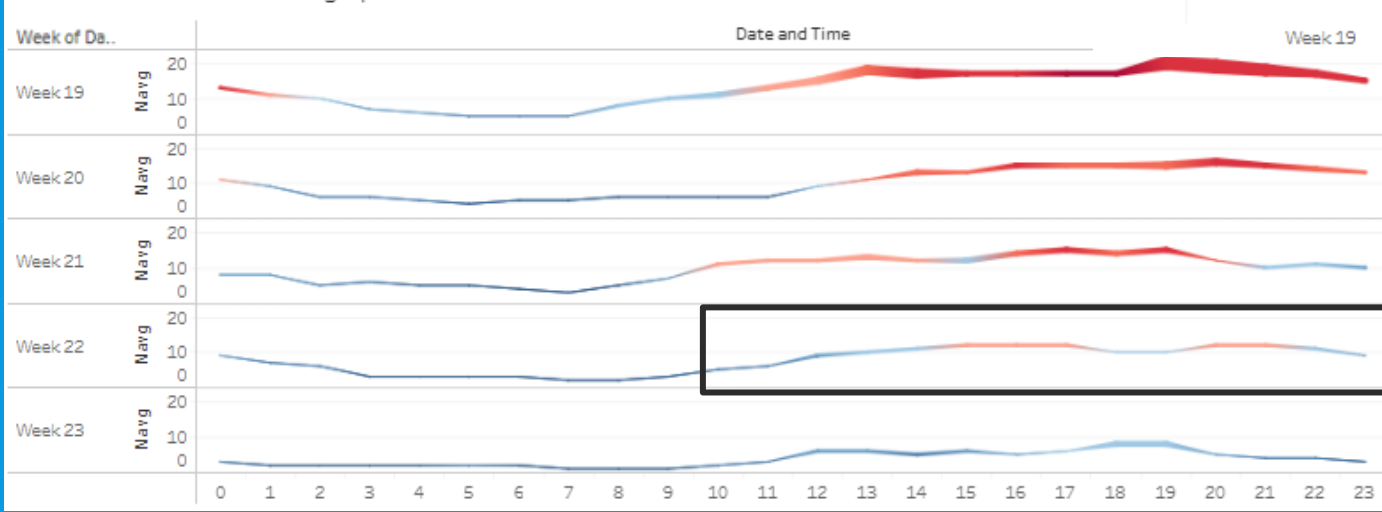
- This chart shows overcrowding occurrences through the month, red (high) and blue (few)
- The NEDOC averages declining week over week.
- This identifies that something was initiated at this ED to mitigate overcrowding and it is working.

Decline of Overcrowding week

Decline of Overcrowding week



Decline of Overcrowding split week



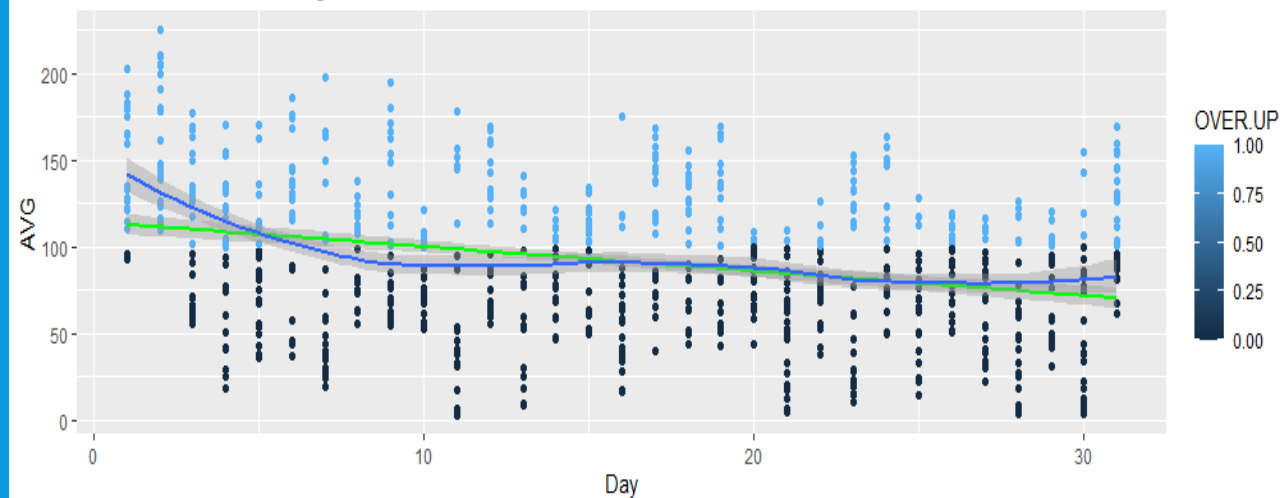
- The second chart shows the same information broken out into individual weeks.
- As of the 4<sup>th</sup> week there is barely any red on the chart.



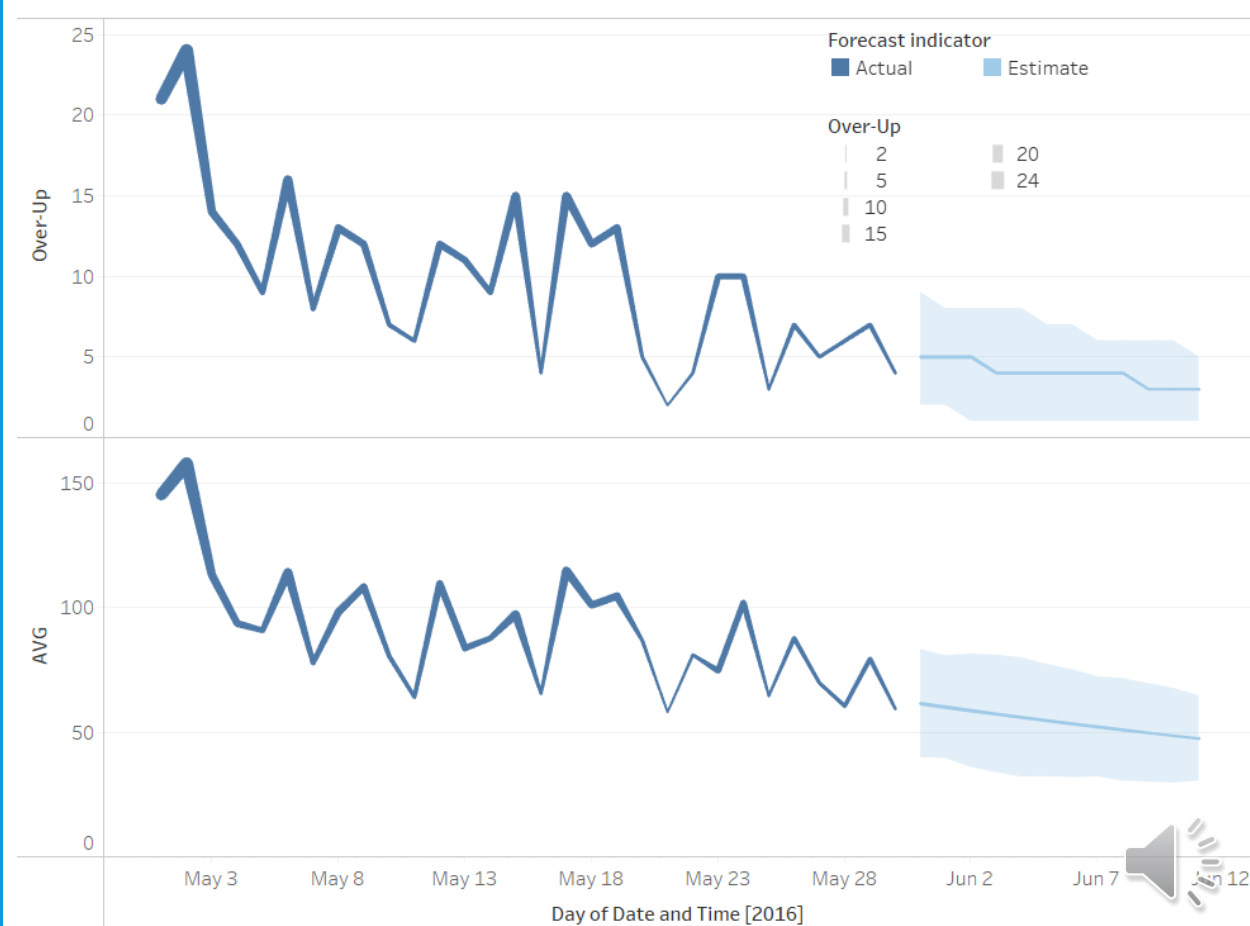
# GENERAL PREDICTIVE ANALYSIS

- The first chart represents 2 different trend lines indicating the downward progression of the NEDOC Average score
- The second forecasts future averages and appears that the downtrend theory has validity.

Decline of Overcrowding



NEDOC Predictive



# SUMMARY OF VISUAL ANALYSIS

- Number of Patients (DEP) and time are extremely significant variables
- *OVR*/overcrowding variable provided excellent trend analysis information.
- Strong correlations exist between patient arrival times, and time spent in the ED.
- NEDOC Scores 0-3 are evenly distributed throughout the data set
  - Disaster or NEDOC Score 4, was not well distributed and was treated as an outlier
- The 5<sup>th</sup> week is not a complete week
- Based on data trends, it can be assumed that the ED implemented something within the month that greatly improved their NEDOC score.



# REVIEW OF PREDICTIVE MODELS

- SVM or Support Vector Machine
  - Compares all variables equally
  - Limited configurability
- Naïve Bayes
  - Needs specific variables identified for the analysis
  - Bases factors on preidentified assumptions.
    - These Pre-predicted outcomes are then used as a basis for overall predictions
- Random Forrest completely randomizes data
  - Compares all variables equally
  - Can include, or exclude variables or even rows of data



# PREDICTIVE ANALYSIS

SVM	0	1	0.13004484
0	114	13	0.8976378
1	16	80	0.83333333
	0.876923	0.860215	0.86995516

Accuracy 87%

NAÏVE	0	1	0.225225
0	86	44	0.661538
1	6	86	0.934783
	0.934783	0.661538	0.774775

Accuracy 77% Threshold @ 75%

LR	0	1	0.15695067
0	117	22	0.841727
1	13	71	0.845238
	0.9	0.763441	0.843049

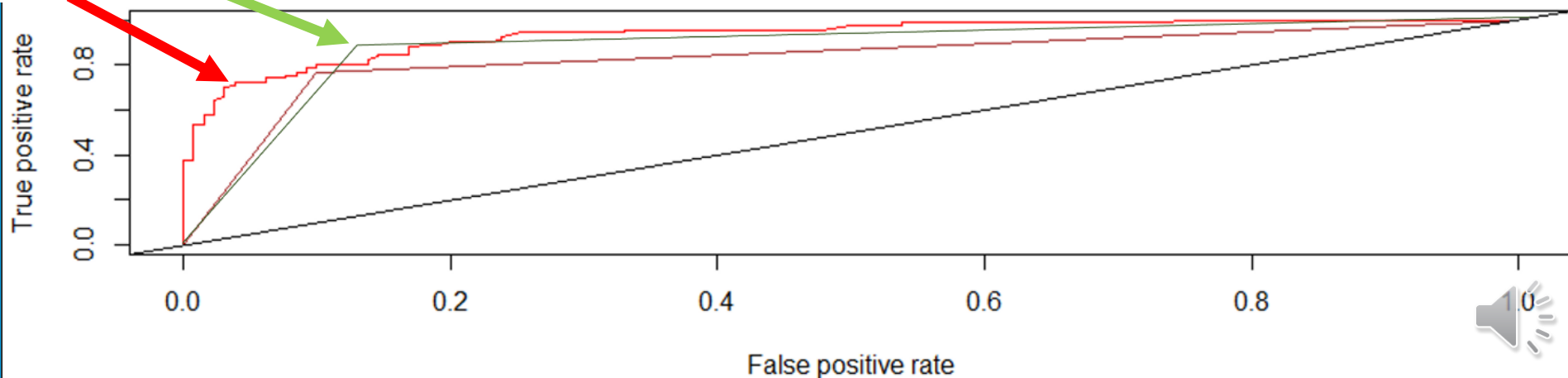
Accuracy 84%

RF	0	1	0.14798206
0	107	23	0.823077
1	10	83	0.892473
	0.91453	0.783019	0.852018

Accuracy 85%

- SVM
  - Highest overall general accuracy and total averages of all models
  - Lowest negativity predictions (false predictions)
- Random Forrest was a close second
- SVM Green on the ROC and the Random Forest Red

ROC Curves



Having these 4 models built from the exact same test and train values will prove beneficial for new data and future data-based inquiries





# MITIGATION OF OVERCROWDING

- Would it be beneficial for the ED to hire more physicians to work peak hours?



# MITIGATION OF OVERCROWDING

- Does the downtime of physicians or admins contribute to overcrowding?





# MITIGATION OF OVERCROWDING

Could the ED divert patients to a neighboring hospital when it gets overcrowded?



# MITIGATION OF OVERCROWDING

Could adding additional hospital beds in the ED be a solution?



# MITIGATION OF OVERCROWDING

Could introducing sperate check-in and discharge locations, help?





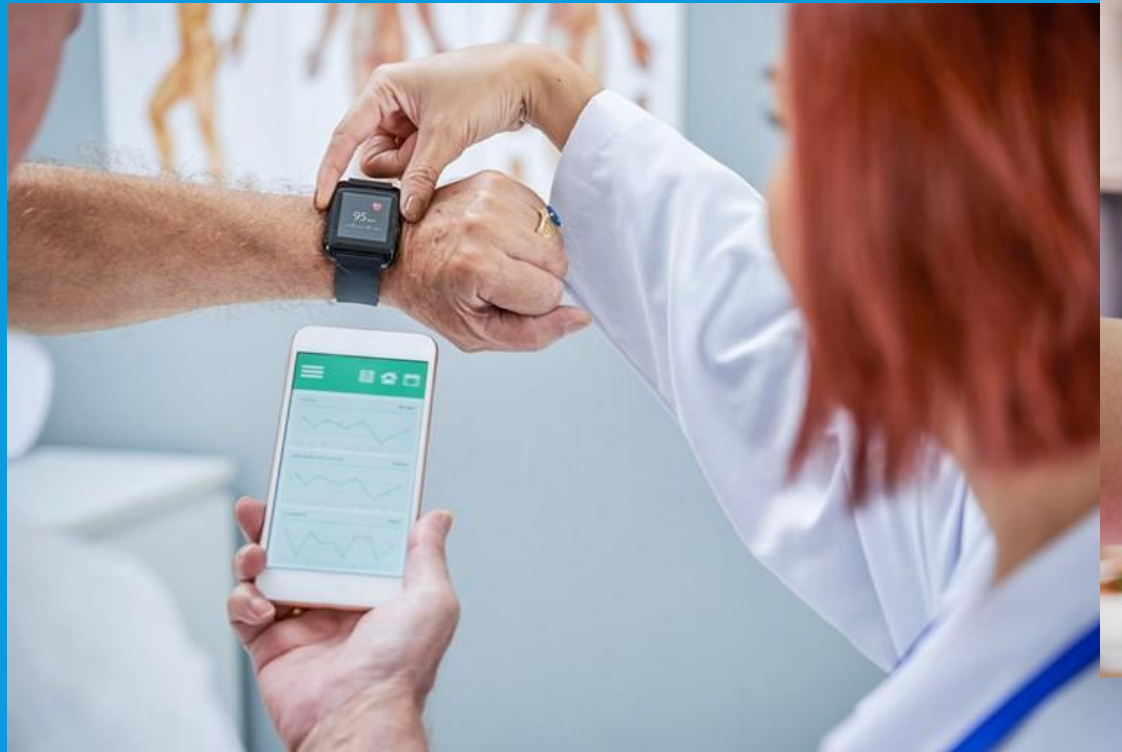
# MITIGATION OF OVERCROWDING

New physicians and trainee's during non-peak hours.



# MITIGATION OF OVERCROWDING

Introduction of wearable technologies for tracking purposes.





# FINAL STATEMENTS

- All variables were consistent, except for the NEDOC scores.
- The Disastrous scores were treated like outliers
- Remaining scores show progression toward more manageable scores.
- NEDOC scores are forecasted to continuing decreasing.

Week	Sum of weekly NAVG	Count Below-Overcrowding	Count Above-Overcrowding	% <u>per</u> week of Overcrowding	% <u>difference</u> in Overcrowding
1	311	64	104	0.62	
2	239	98	70	0.42	-0.33
3	229	102	66	0.39	-0.06
4	182	123	45	0.27	-0.32
5	89	46	26	0.36	-0.42

- The ED is well on its way to fixing the Overcrowding issue!

