

TS Viewership Analysis

Vincent Messina

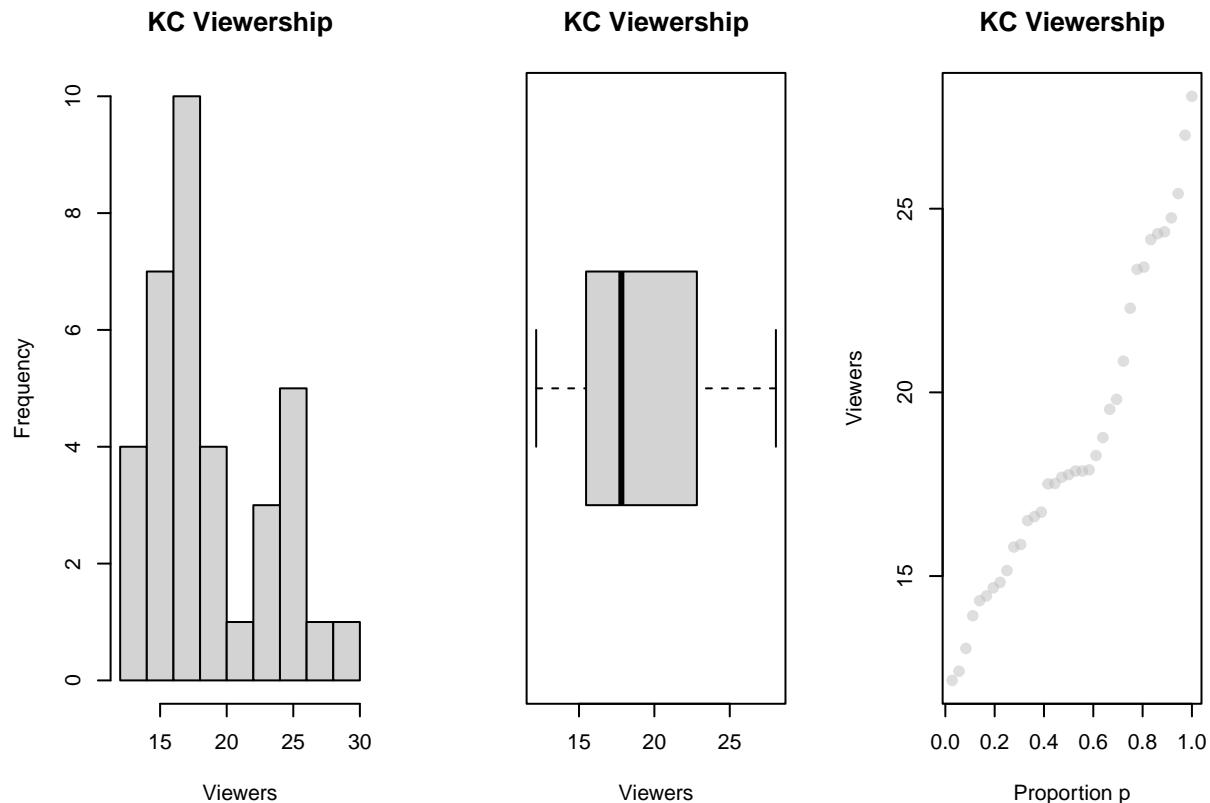
2023-11-28

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	12.16	15.63	17.81	18.75	22.55	28.06

Opponent Viewers

## 1	ARIZ	16.6200
## 2	BAL	19.8100
## 3	BUF	21.4650
## 4	CHI	24.3200
## 5	CIN	21.0900
## 6	CLE	19.5400
## 7	DAL	28.0600
## 8	DEN	17.5100
## 9	DET	24.7500
## 10	GB	24.3700
## 11	HOU	14.8300
## 12	IND	14.6800
## 13	JAX	15.4850
## 14	LAC	15.7775
## 15	LAR	23.3500
## 16	LV	16.0620
## 17	NYG	13.9200
## 18	NYJ	27.0000
## 19	PHI	18.2800
## 20	PIT	24.1600
## 21	SEA	15.8600
## 22	SF	22.2900
## 23	TB	20.8500
## 24	TEN	16.4200
## 25	WFT	12.4100

The match-up against DAL appears to be the most exciting as it had the highest average viewership
while the match-up against WFT appears to be the least exciting based on the viewership data.



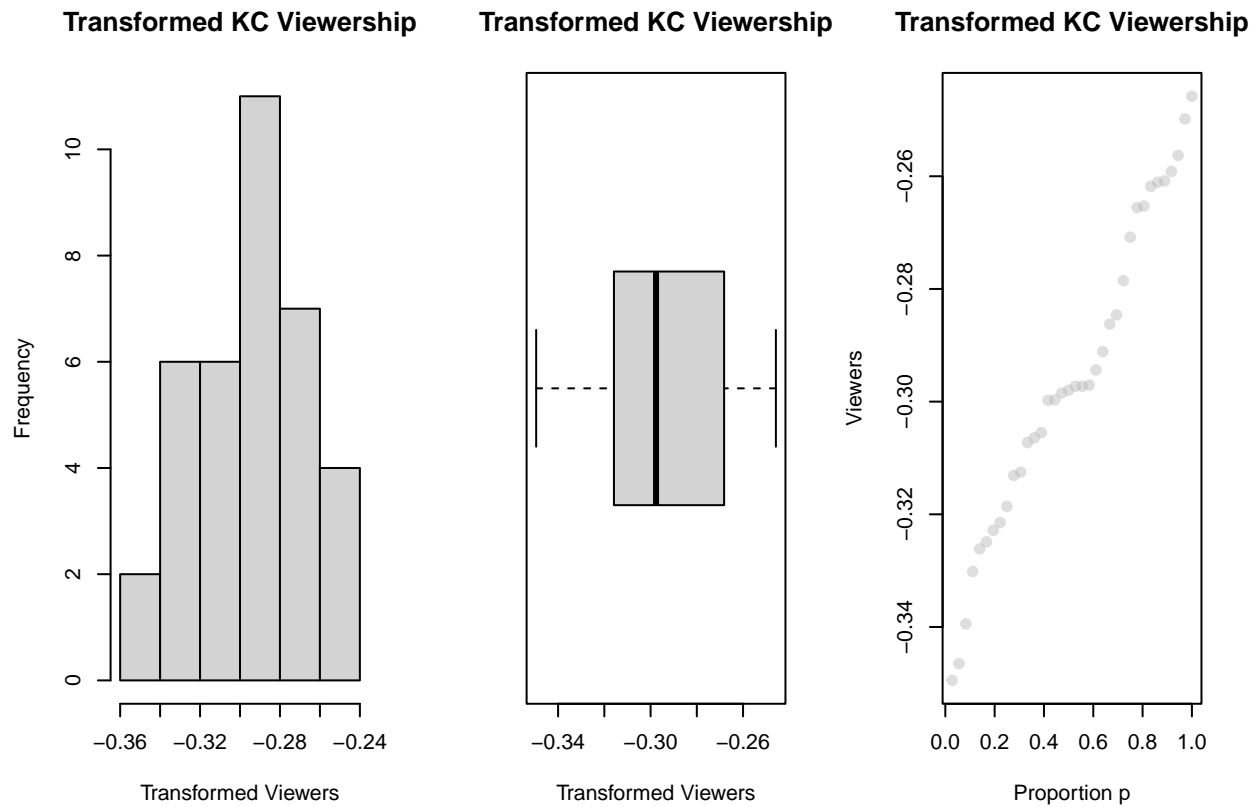
Histogram: The distribution of viewers appears somewhat right-skewed, meaning there are a few weeks with exceptionally high viewership numbers compared to the rest. Most weeks fall within the 15 to 25 million viewers range, with a peak around 20 million.

Boxplot: The median viewership seems to be around 18 million. The interquartile range suggests that 50% of the weeks have viewership numbers between approximately 15 million and 23 million. There are several outliers on the higher end, which aligns with the right-skew observed in the histogram.

Quantile Plot: The points deviate from the straight line, especially on the right end. This suggests that the viewership distribution doesn't strictly follow a normal distribution. The rightward bend at the high end of the Quantile-plot further confirms the right-skew observed in the histogram.

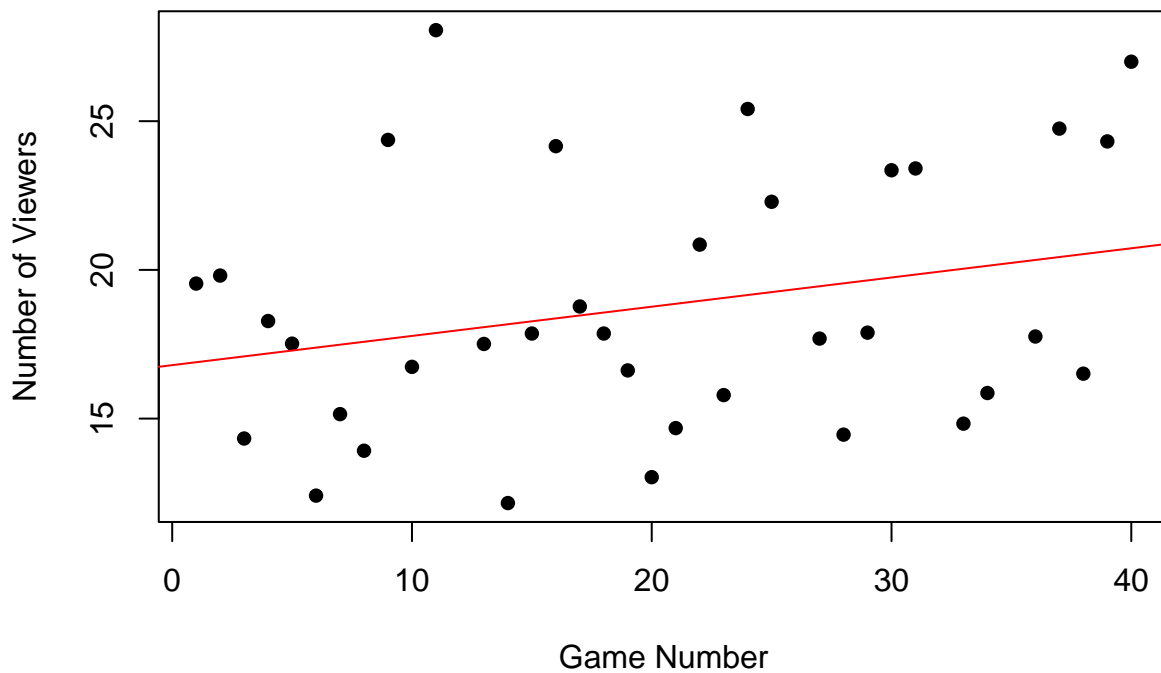
Based on the above observations: The "typical" range for weekly viewership might be considered around 15 million to 23 million. Weeks with viewership outside this range are atypical. Likely when Taylor Swift attends.

```
## [1] -0.4208375
```

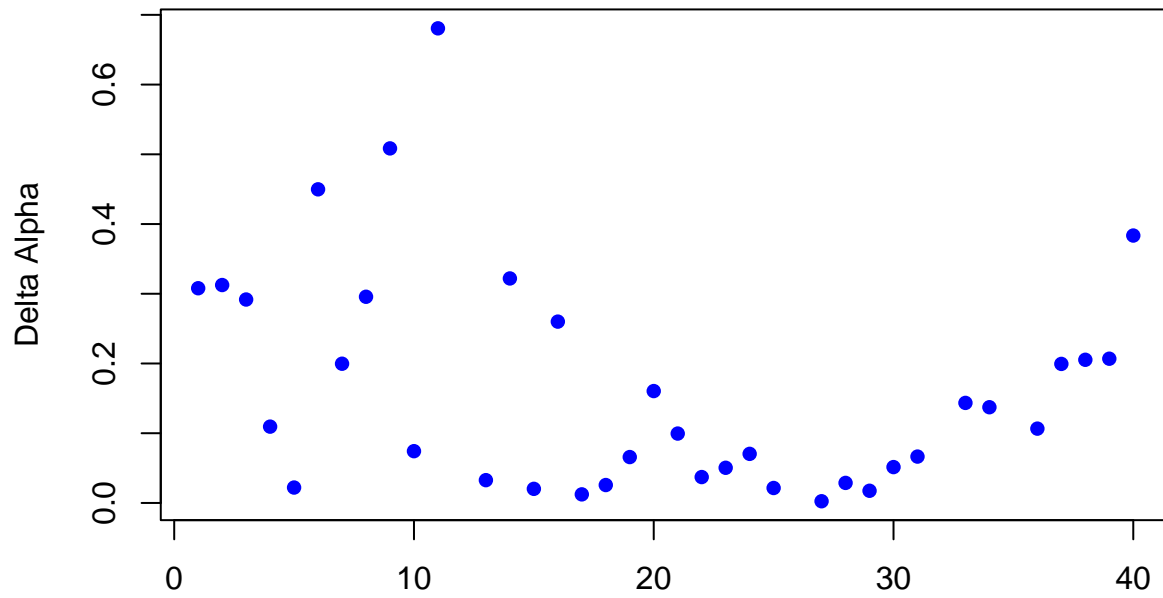


The data appears to have slight left skewness. The bump rule from class could be preferable because we could fine tune the value to the given data.

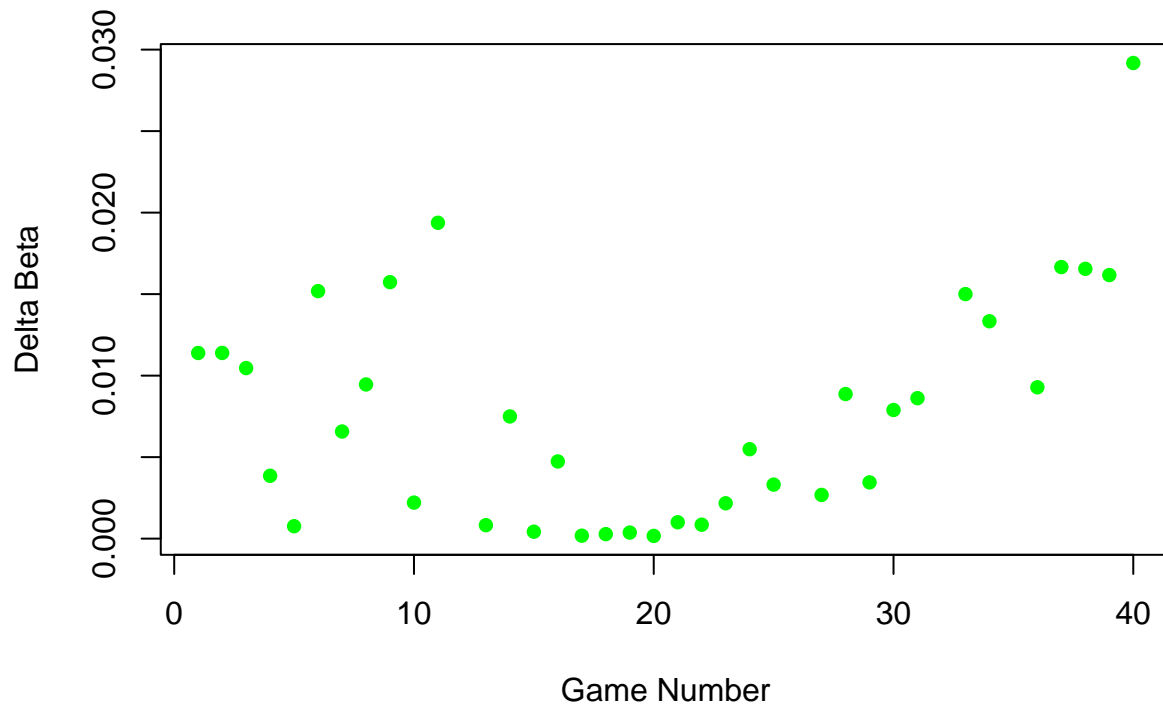
Scatter Plot of Viewers vs Game_Num



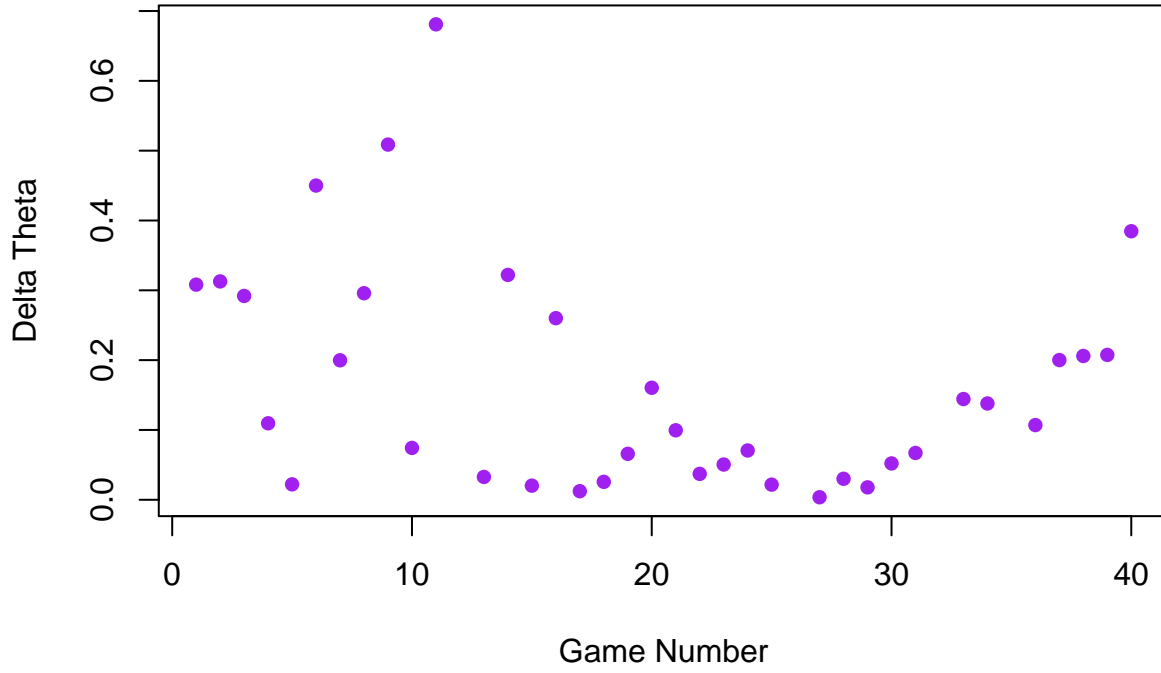
Influence Plot for Alpha



Game Number
Influence Plot for Beta

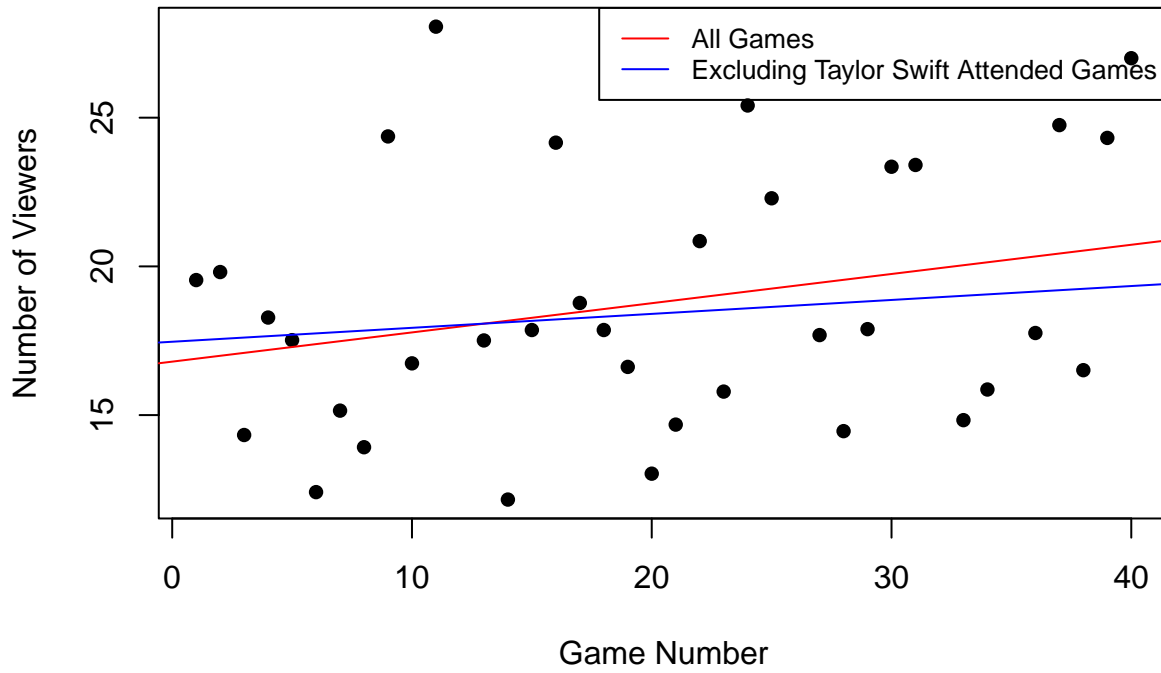


Influence Plot for Theta



Top 4 influential games: 26 28 31 1

Scatter Plot of Viewers vs Game_Num



Given the Tukey's bisquare (or biweight) objective function:

$$\rho_k(r) = \begin{cases} \frac{r^2}{2} - \frac{r^4}{2k^2} + \frac{r^6}{6k^4} & \text{for } |r| \leq k \\ \frac{k^2}{6} & \text{for } |r| > k \end{cases}$$

For $|r| \leq k$: The derivative with respect to r is:

$$\frac{\partial}{\partial r} \left(\frac{r^2}{2} - \frac{r^4}{2k^2} + \frac{r^6}{6k^4} \right) = r - \frac{2r^3}{k^2} + \frac{r^5}{k^4}$$

For $|r| > k$: The derivative with respect to r is:

$$\frac{\partial}{\partial r} \left(\frac{k^2}{6} \right) = 0$$

Given $r_u = y_u - \alpha - \beta x_u$, the gradient can be calculated as:

For $|r| \leq k$:

$$g_1 = \frac{\partial \rho}{\partial \alpha} = \sum_{u \in \mathcal{P}} \left(-1 \times \left(r - \frac{2r^3}{k^2} + \frac{r^5}{k^4} \right) \right)$$

$$g_2 = \frac{\partial \rho}{\partial \beta} = \sum_{u \in \mathcal{P}} \left(-x_u \times \left(r - \frac{2r^3}{k^2} + \frac{r^5}{k^4} \right) \right)$$

For $|r| > k$:

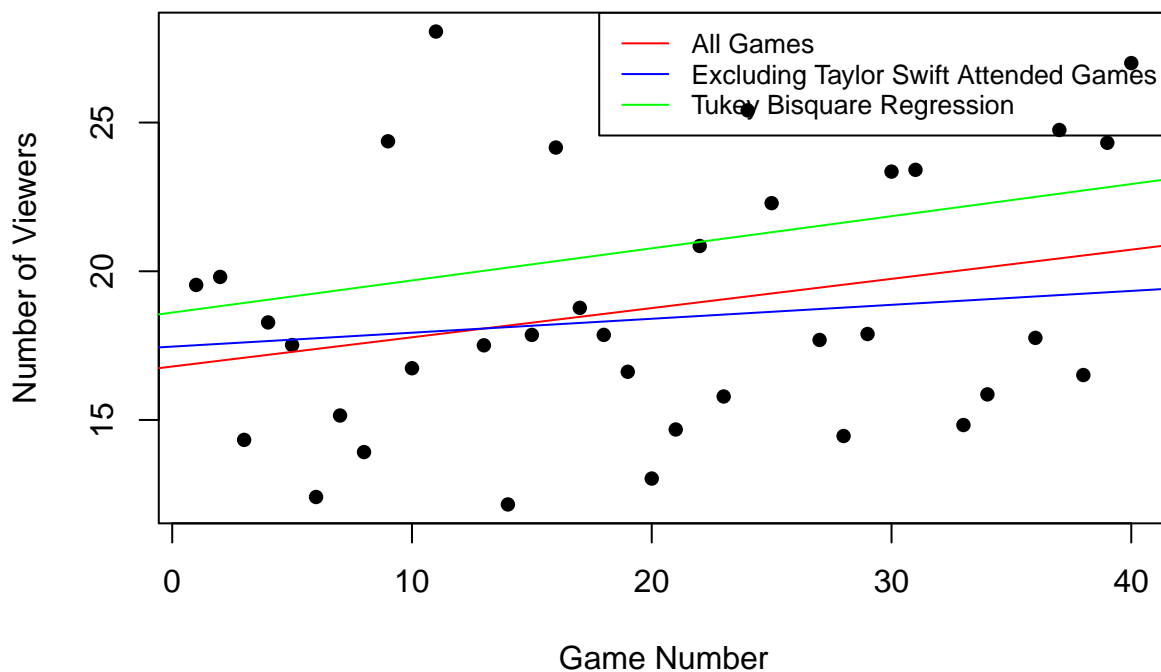
$$g_1 = g_2 = 0$$

Thus, the gradient vector for $|r| \leq k$ is:

$$\mathbf{g} = \nabla \rho(\boldsymbol{\theta}; \mathcal{P}) = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

```
## [1] "Optimized parameters from gradient descent:"
## $theta
## (Intercept)    Game_Num
## 18.6111648    0.1080424
##
## $converged
## [1] TRUE
##
## $iteration
## [1] 270
##
## $fnValue
## [1] 280.1348
##
## [1] "Parameters from robust regression:"
## (Intercept)    Game_Num
## 16.5403610    0.1056097
```

Scatter Plot of Viewers vs Game_Num



The Tukey seems to do the best job, it seems to be the most centred and capture everything the best. However there isn't a very big difference. You could see the blue line is the most flat. The viewership hasn't really changed across the season. You notice that when you include Taylor (the red line) there is actually an increase in the number of viewers as the season goes by, so with only 2 games affected, she certainly had a noticeable influence.