SNL plots

Fred & Ella

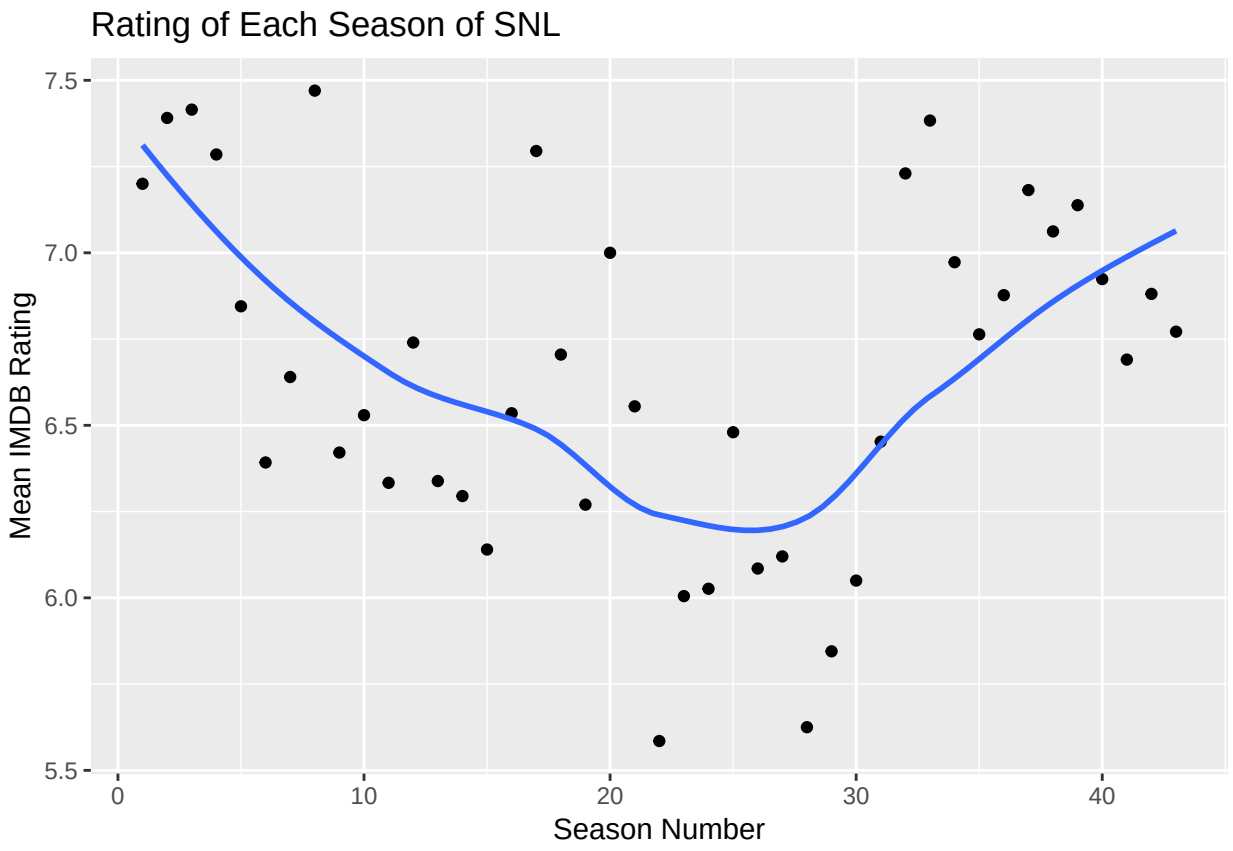
11/23/2019

SNL DATA and Why?

After gathering data from wikipedia and IMDB on Saturday Night Live episodes, hosts, and ratings, we are trying to discover how the perception of Saturday Night Live has changed over time. This relates to ratings and viewership. We are also seeking to find out if bringing in a popular host can positively impact the ratings. Likewise, SNL has a five timers club which is composed of any hosts who have hosted at least five times. We will be analyzing this group as well to see if there is a positive correlation between the ratings of the episode and the hosts who are in the five timers club.

Looking first at the mean IMDB ratings for each season, we compiled all of the ratings per episode within a season and took their average, to find a general trend for how ratings have changed as time has progressed.

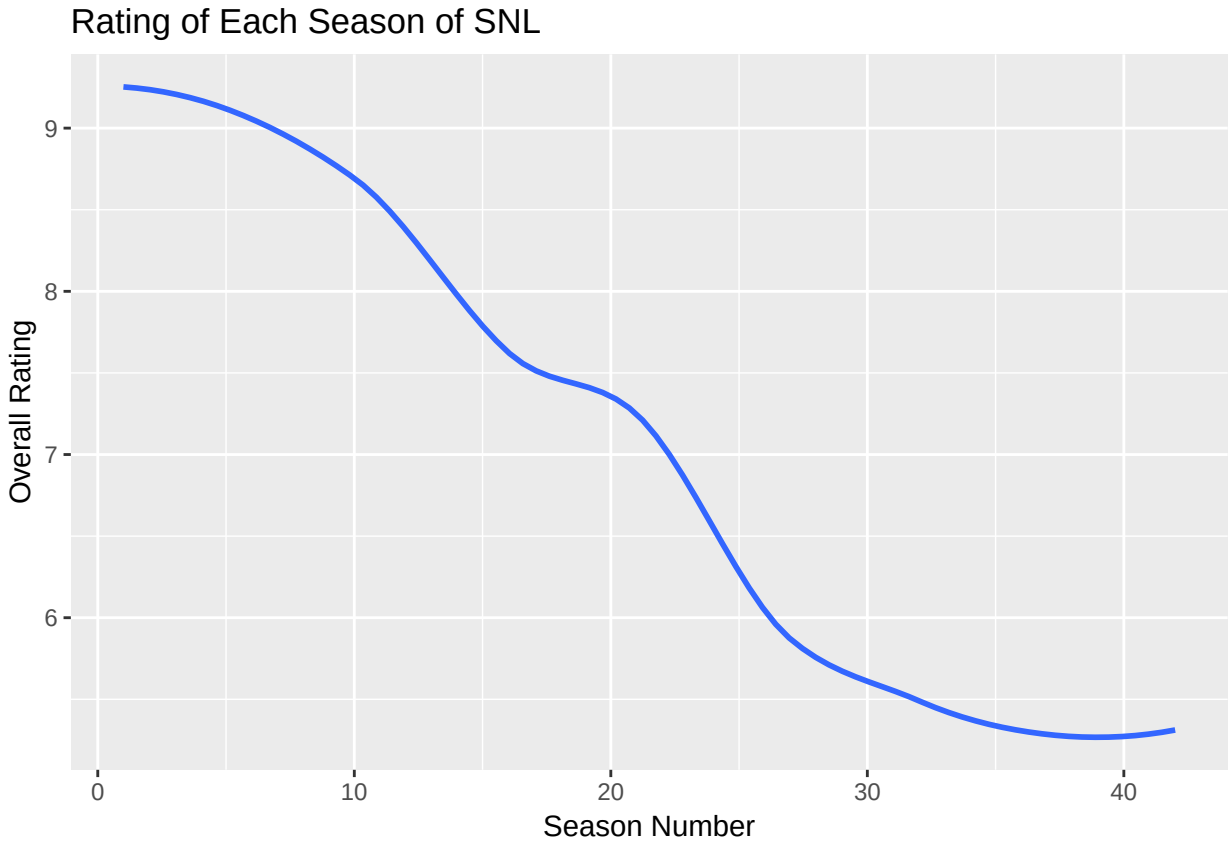
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



The IMDB ratings show that there was a dip in ratings around season 20, but ratings increased again after season 30. Ratings began quite high, and ended high, but in the middle there was a substantial dip in ratings. This shows in the IMDB ratings, but what do other ratings say?

Looking now at the overall rating for each season is compiled from the TV Ratings Guide. While this data is incomplete for all seasons in the data set, ideally it will still be able to provide some insight into the ways that ratings have changed over time.

```
## Warning: 21 parsing failures.
## row col expected actual
## 830  -- a double    N/A
## 831  -- a double    N/A
## 832  -- a double    N/A
## 833  -- a double    N/A
## 834  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 21 parsing failures.
## row col expected actual
## 830  -- a double    N/A
## 831  -- a double    N/A
## 832  -- a double    N/A
## 833  -- a double    N/A
## 834  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 21 rows containing non-finite values (stat_smooth).
```



This plot shows information contrary to the IMDB ratings. Rather than a recent increase in rating, it shows a steady decrease over time since the initial seasons of Saturday Night Live. In fact it shows an all time low reached around season 40.

Perhaps the hosts or musical guests impact the ratings. Here we have lists of the best and worst rated hosts

and musical guests.

```
## # A tibble: 576 x 3
##    Host              mean_IMDB n_hosts
##    <chr>                 <dbl>   <int>
##  1 Betty White            8.8        1
##  2 Richard Pryor          8.6        1
##  3 Justin Timberlake      8.30       5
##  4 Maya Rudolph           8.2        1
##  5 Bill Hader             8.15       2
##  6 Kristen Stewart        8.1        1
##  7 Ron Howard             8.1        1
##  8 William Shatner        8.1        1
##  9 Andy Samberg           8          1
## 10 Ariana Grande          8          1
## # ... with 566 more rows
```

```
## # A tibble: 576 x 3
##    Host             mean_IMDB n_hosts
##    <chr>                <dbl>   <int>
##  1 Paris Hilton          3.8        1
##  2 Al Sharpton           3.9        1
##  3 Cuba Gooding Jr.      3.9        1
##  4 Pamela Anderson       4          1
##  5 Rosie O'Donnell       4.15       2
##  6 Donald Trump          4.3        2
##  7 Justin Bieber         4.3        1
##  8 Nia Vardalos          4.7        1
##  9 Bill Pullman          4.8        1
## 10 January Jones         4.8        1
## # ... with 566 more rows
```

Here we see that there seems to be no correlation between the ratings the host recieves on the show and the number of times they host.

```
## # A tibble: 614 x 3
##    Musical_Guest                            mean_IMDB n_Musical
##    <chr>                                        <dbl>     <int>
##  1 Gil Scott-Heron                                8.6         1
##  2 Queen                                          8.5         1
##  3 Justin Timberlake                              8.4         4
##  4 Soundgarden                                    8.4         1
##  5 The Blues Brothers                             8.3         2
##  6 Carly Simon                                    8.2         1
##  7 Ciara                                          8.2         1
##  8 Sleigh Bells                                   8.2         1
##  9 The 1975                                       8.2         1
## 10 3-D, Paul McCartney and Linda McCartney        8.1         1
## # ... with 604 more rows
```

```
## # A tibble: 614 x 3
##    Musical_Guest            mean_IMDB n_Musical
##    <chr>                        <dbl>     <int>
##  1 Keane                         3.8          1
##  2 Rollins Band                  4            1
##  3 James TaylorDon Grolnick      4.2          1
```

```
##  4 Ricky Martin                  4.4          2
##  5 Ms. Dynamite                  4.5          1
##  6 50 Cent                       4.75         2
##  7 Dwight Yoakam                 4.8          1
##  8 New Edition                   4.8          1
##  9 Pink                          4.85         2
## 10 Aretha Franklin               4.9          1
## # ... with 604 more rows
```

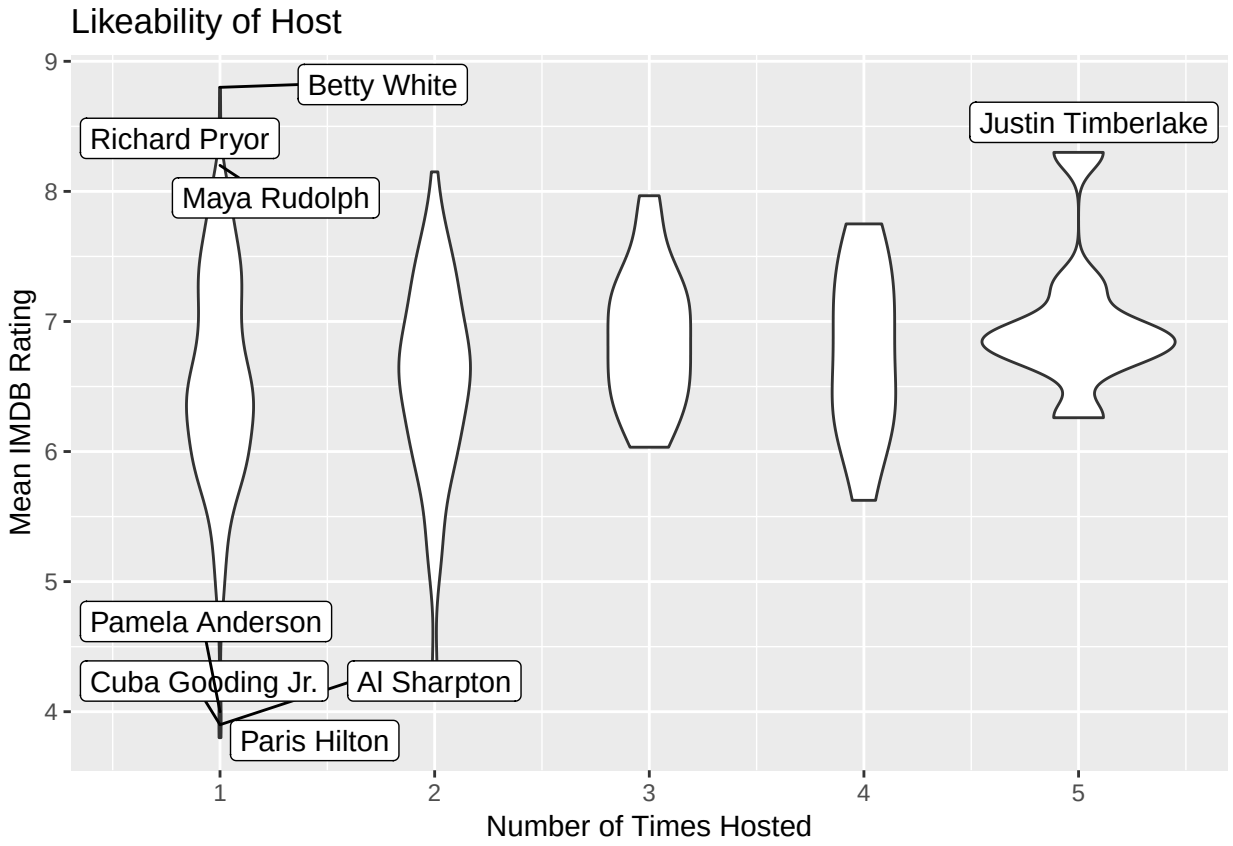
Here we see again the number of times a musical guest is invited back to host is not dependent on the rating the episode recieves.

So who is invited back to the show and what is their mean rating?

```
## # A tibble: 576 x 3
##    Host               mean_IMDB n_hosts
##    <chr>                  <dbl>   <int>
##  1 Alec Baldwin            6.74      16
##  2 Steve Martin            7.69      14
##  3 John Goodman            6.51      13
##  4 Buck Henry              7.52      10
##  5 Tom Hanks               6.6        9
##  6 Chevy Chase             6.91       7
##  7 Christopher Walken      7.34       7
##  8 Drew Barrymore          6.7        6
##  9 Elliott Gould           6.87       6
## 10 Ben Affleck             6.78       5
## # ... with 566 more rows
```

```
## # A tibble: 614 x 3
##    Musical_Guest mean_IMDB n_Musical
##    <chr>             <dbl>     <int>
##  1 Beck               6.51         7
##  2 Foo Fighters       6.51         7
##  3 Kanye West         6.77         6
##  4 Paul Simon         6.73         6
##  5 Coldplay           6.92         5
##  6 Eminem             6.56         5
##  7 Maroon 5           6.9          5
##  8 Rihanna            6.98         5
##  9 Sting              6.92         5
## 10 Aerosmith          6.35         4
## # ... with 604 more rows
```

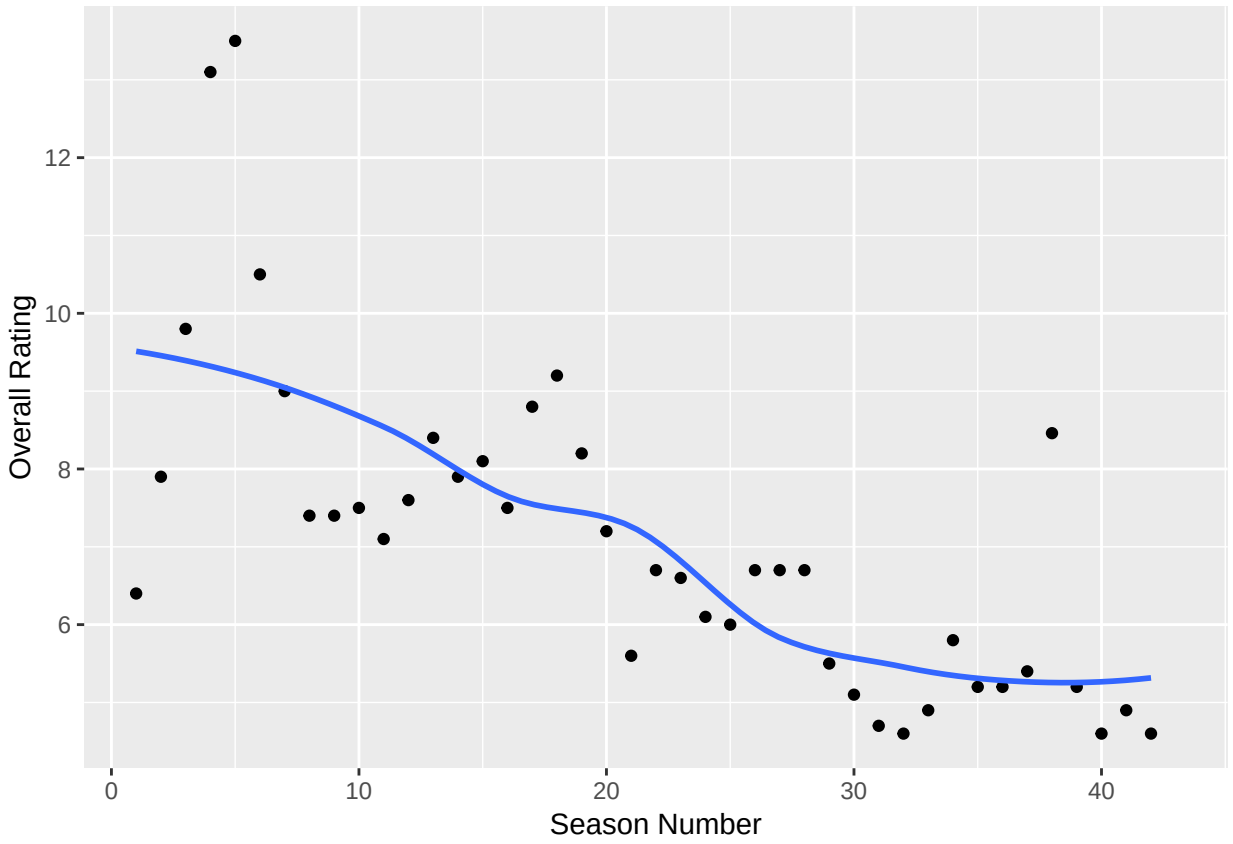
Here we see that the Host and Musical Guest that are invited back to the show have a mean IMDB rating between 6 - 7.5. Here we see that hosts and Musical Guests who are invited back to the show are people with more name recognition.



Here plotting how many times a host is invited back we see that hosts who have only hosted one time has the greatest spread in IMDB Ratings, while the more times you are asked to host the higher mean IMDB Rating.

These next plots show overall rating, 18-49 rating, and viewers in millions by season. Each of these plots show a decrease as time progresses. Especially noting that seasons 31-33 had very poor reviews and viewership. While it has improved a little since then, the reviews and quantity of viewers is nowhere near what it had been initially.

```
## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A

## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A

## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```



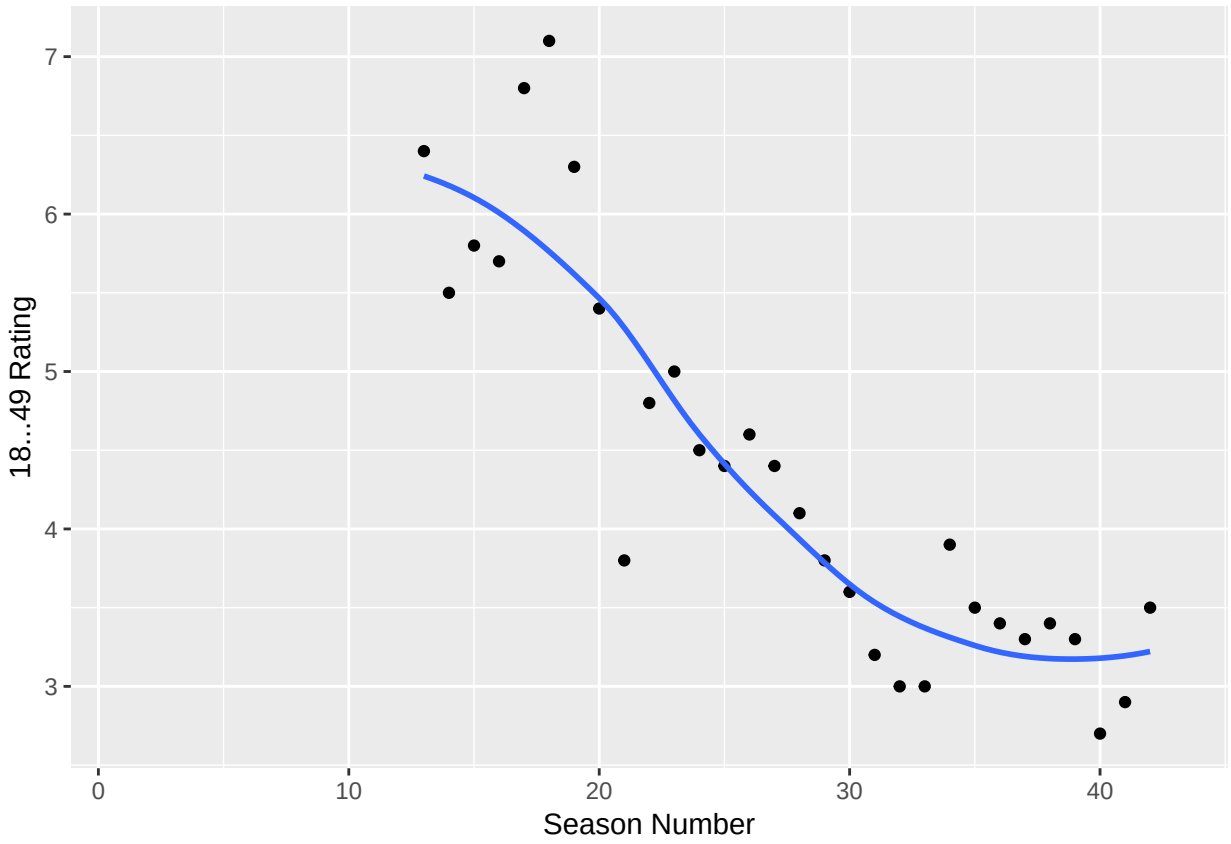

```
## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
```

```
## ... ... ........ ......
## See problems(...) for more details.

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 13 rows containing non-finite values (stat_smooth).

## Warning: Removed 13 rows containing missing values (geom_point).

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
```

```
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label),
## x$x, x$y, : conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot
## substituted for <e2>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label),
## x$x, x$y, : conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot
## substituted for <80>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label),
## x$x, x$y, : conversion failure on '18-49 Rating' in 'mbcsToSbcs': dot
## substituted for <93>
```




```
## Warning: 12 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 12 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

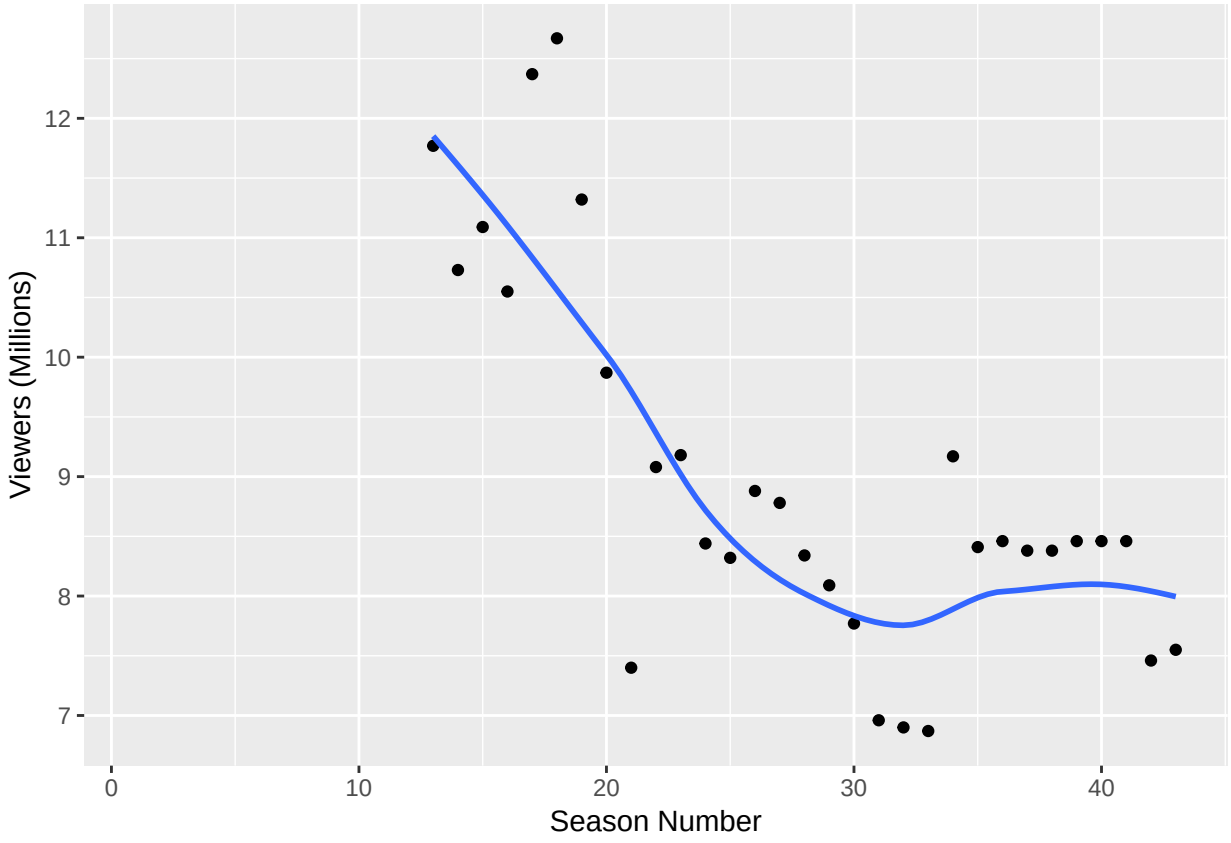
```
## Warning: 12 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
```

```
## ... ... ........ ......
## See problems(...) for more details.
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 12 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 12 rows containing missing values (geom_point).
```




```
## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A
```

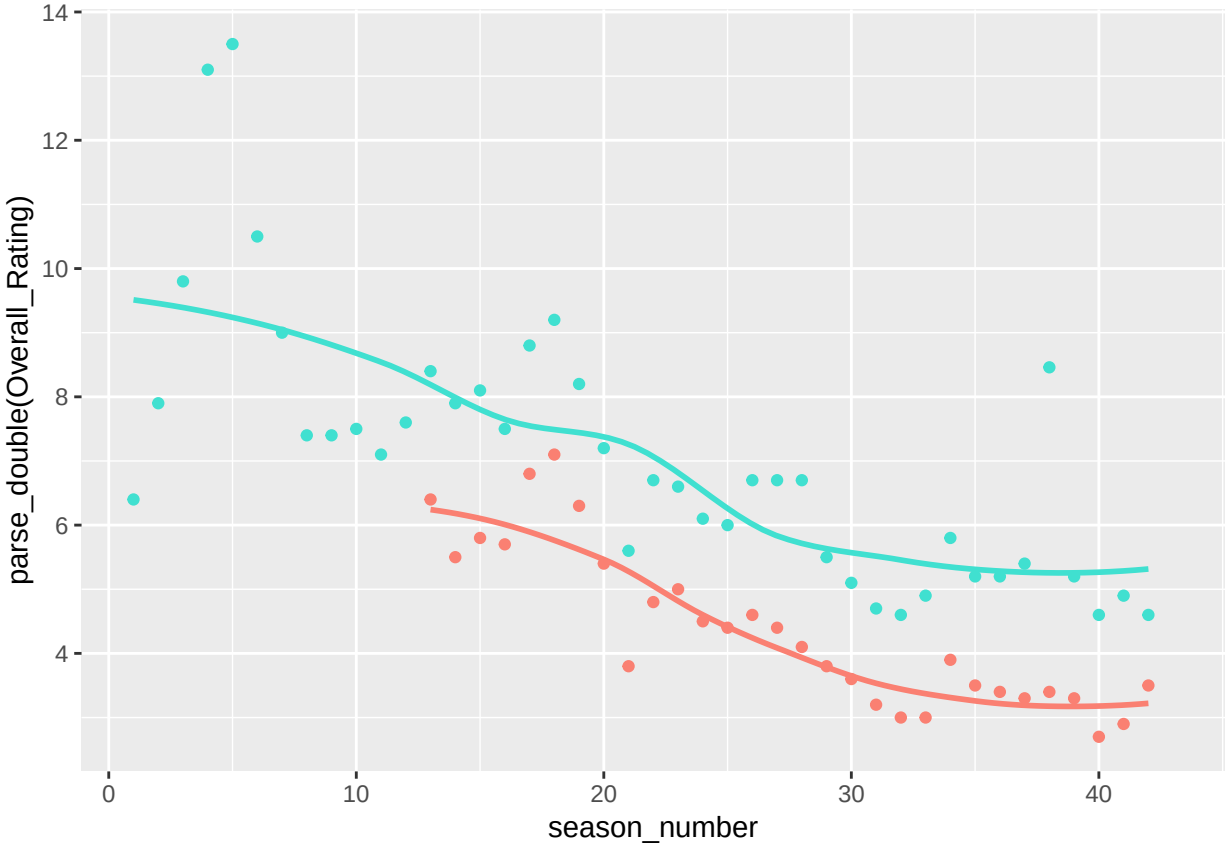
```
## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A
```

```
## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 1 parsing failure.
## row col expected actual
##  43  -- a double    N/A
```

```
## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
## Warning: 21 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 24 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
```

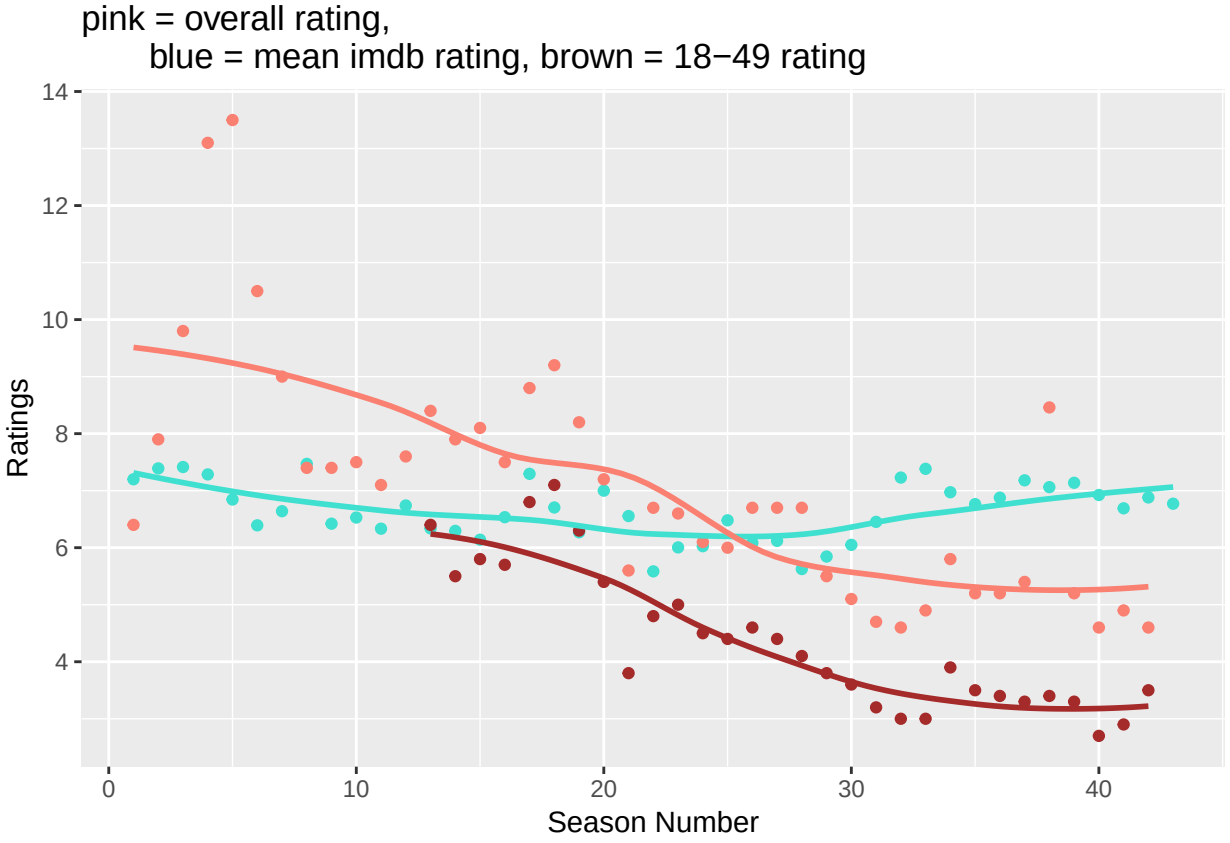
```
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 22 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 13 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
```

```
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 19 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 17 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 18 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.

## Warning: 20 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
```

```
## ... ... ........ ......
## See problems(...) for more details.
```

```
## Warning: 21 parsing failures.
## row col expected actual
##   1  -- a double    N/A
##   2  -- a double    N/A
##   3  -- a double    N/A
##   4  -- a double    N/A
##   5  -- a double    N/A
## ... ... ........ ......
## See problems(...) for more details.
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



This last plot shows the changes in mean IMDB rating, overall rating from TV Ratings Guide, and the 18-49 rating which is the rating from those aged 18 to 49 years old as the season changes. IMDB ratings are shows in blue, the overall rating is shown in pink, and the 18-49 rating is shown in brown. Both the overall rating and 18-49 rating show a steady decrease as time progresses. Only the mean IMDB rating shows that recent seasons have improved after a decrease.