**Allen Miller 62%**

**EDA: (10 % of overall grade in the class.)**

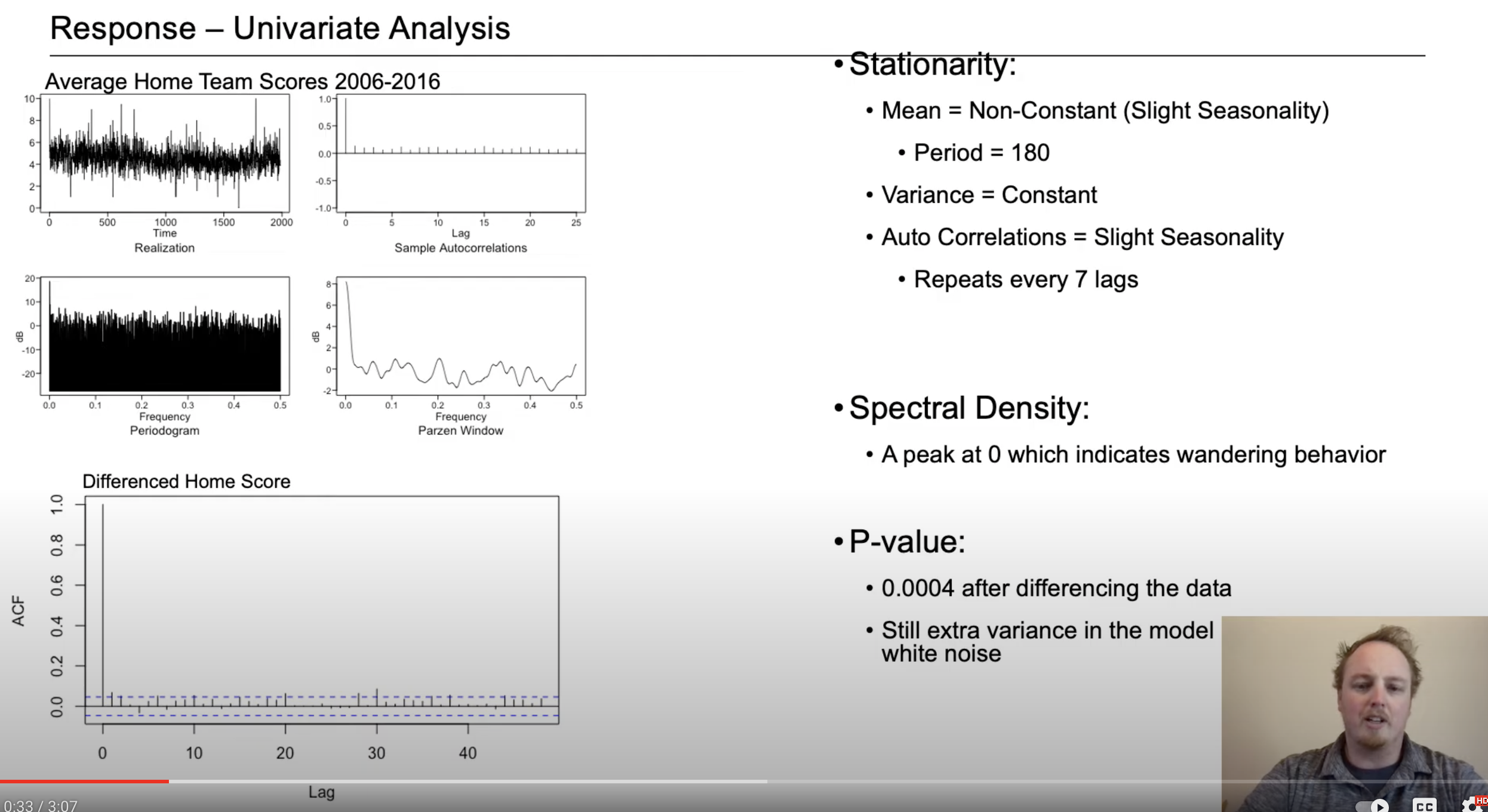
**Sunday July 11th at 11:59pm**

Deliverable:

1. 3-minute YouTube (or Zoom) video:
2. Identify yourself. Good job.
3. Describe Data Set / Time Series (Who, What, When, Where, Why and How)

Good job describing the data although I am not completely sure what the series is. Average scores for all MLB teams? Is this the average score for all teams together or an average score for each team? Also, are the time points evenly spaced? You have 2000 time points for roughly 11 years. Are these daily average scores?

Stationary / Non-Stationary

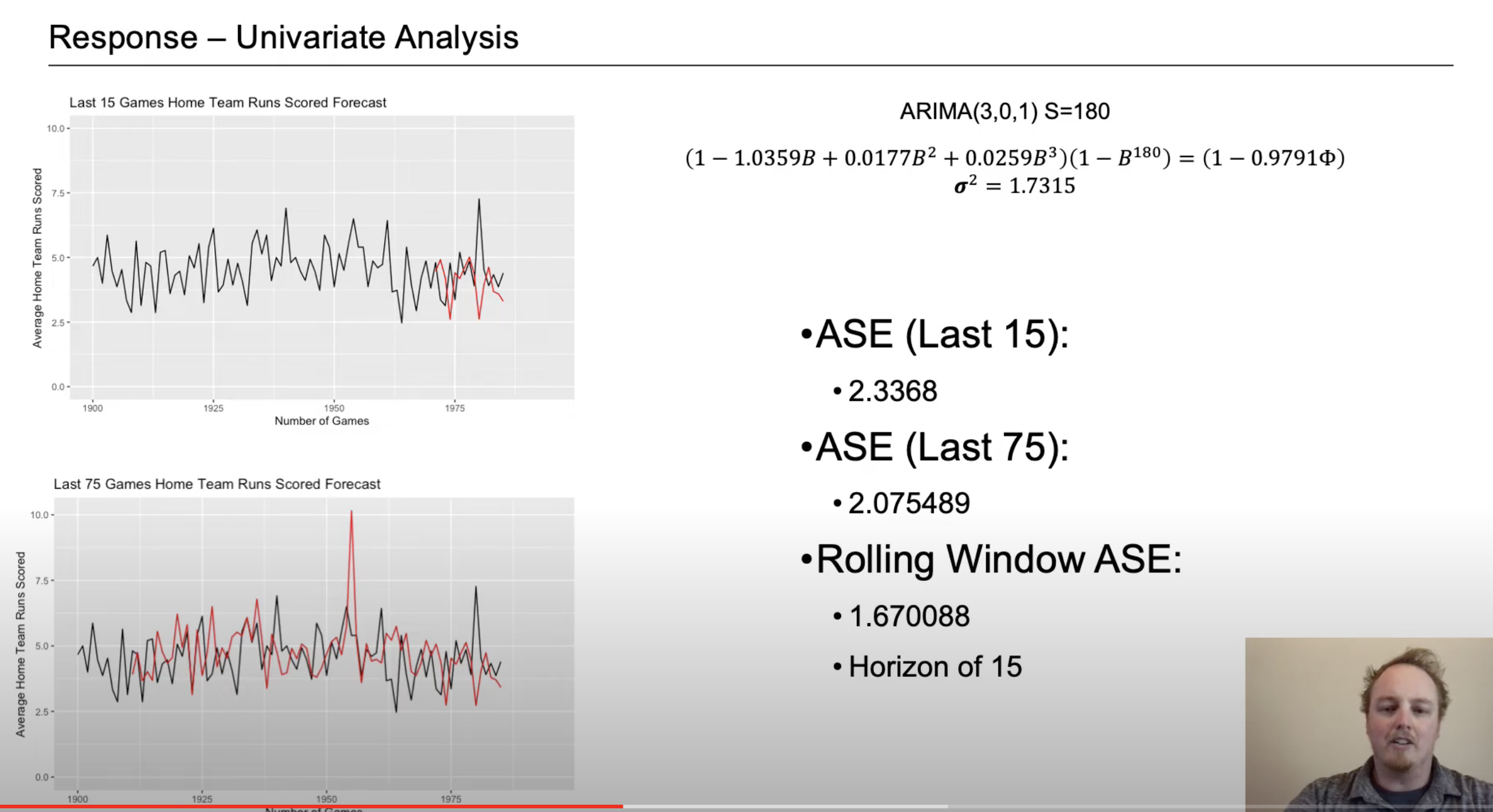


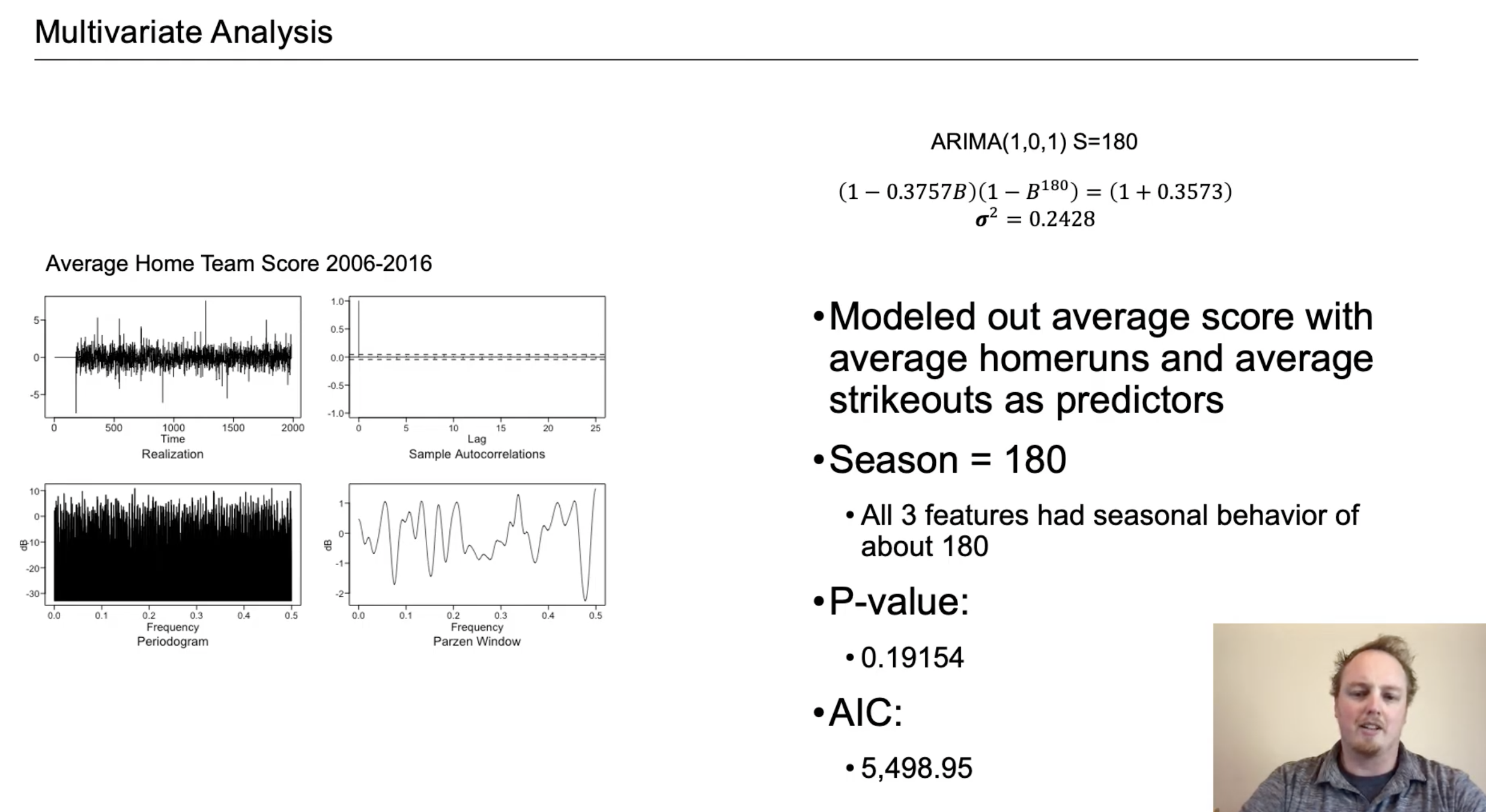
I am still weary of what the data actually is, but this series looks like white noise even before the difference. The ACF in the lower left says Differenced Home Score but is the ACF of the differenced Home Team Score. Also, it is said in the presentation that there is a peak in the spectral density at .6. The highest frequency that can be measured is .5 (the Nyquist Frequency). There really doesn’t look to me like there is much more than a real peak at zero in this spectral density.

1. ACFs and Spectral Densities just to explore.

Got it.

1. At least 2 candidate ARMA / ARIMA models





* 1. The models in factored form or at least separate the stationary and non-

stationary factors with standard deviation or variance of the white noise. Good job including the models, but note that there is no in the left side and no in the right. Also, the phi on the right side needs to be a B.

* 1. AIC Didn’t see AIC
  2. ASE Good work on ASE
  3. Rolling Window ASE Good job on this!
  4. Visualization of Forecasts for both the short- and long-term Horizons.

1. Strategy / Plans for the rest of the analysis. Good work!
2. Submit your slides to 2DS Good!
3. Submit Rmd file or Jupyter Notebook Didn’t see this.
4. Submit Knit Rmd File or Jupyter Notebook as a pdf, html or docx. Didn’t see this?
5. Make sure your video URL is on the Google Doc. Good!

**Rubric for both EDA**

Knit RMD: 20%

Well organized and documented RMD knit to pdf or html or docx.

0% … you can still submit this.

Results / Analysis: 40%

Correct Interpretation

Creating Useful Models

Performing a Complete Analysis: Model ID, Model Building, Forecasting, Cross Validation

30%

Presentation: 40%

* Communication and presentation of your findings are critical to being a successful data scientist. You will be graded on:
  + Voice inflection
  + Slide Organization / Content
  + Visualization
  + Animation
  + Composure: This will include **not reading** off of the slides and smoothness of delivery.
  + Pace: Not going a second over time. (3 min for EDA and 7 min for Final Presentation.) Your client is very strict on this point.

Great job on time as well.

Great presentation! One of the best!

32%