Script Notes for Introductory Videos

The goal of this document is to help you think in advance about what you want to say in each of the introductory videos. You do not need to deliver these notes verbatim on the day of the shoot. You can use this document to write out exactly what you want to say, or simply list topics, keywords, and phrases. I will include some tips below. In general, try to 'be natural,' and say it how you would say it.

- I have created a section following these tips where you can add notes for the Course Intro, and each weekly introduction.

Jason's Tips for For Course Intro Content:

- Welcome them to the Course. Practice saying the official title of the course. Welcome to…*MBAO550 Course Title* ….
- My name is *[name]* and I will be your professor for the next ten weeks. During that time we will cover *[just a few main topics listed here]* among many other concepts. The goal of this course is to prepare you for…. *[what benefit(s) can they hope to enjoy]*
- Cover some of your background related to the course… I began my work in this area as a *[student, phd candidate, etc.]* at *[university]*. Since then, I have *[give brief related work or teaching history].* I joined Bryant University in [year], where I am a member of the department and teach courses in …
- Make the study of this subject timely and share your joy …. These are exciting times for the study of [topic], as …Current events… interesting facts about the industry…
- Thank you for joining me for this course. I look forward to working with you online.

Jason's Weekly Intro Notes:

As you prepare notes for each weekly introduction, seek to be '*specifically vague*.' For example, instead of "On Wednesday, you will submit a research paper," consider something that conveys the idea but is less detailed. So… "Later this week, you will research [this week's topic] and share your ideas with me." This will help the video to have a longer shelf life, in case you later decide to change the assignment due date, or switch it from a paper to a discussion board.

In fact, rather than focusing on assignments, it's better to speak about the learning concepts that will be covered, and to describe why they are important.

- Begin by saying *'This Week, we will [study, look at, consider, explore] [central topic or learning objective for the week]*. Note… we say this week and not week 1 in case you want to change the order of the course later, the video won't need to be reshot.
- Next cover: Why is this topic important? What will they do with the skills and knowledge they gain this week? Will it help them pass an exam? How will they use these skills in the professional world? *You don't have to answer all of these questions. These are merely prompts to help you generate content.*
- What kinds of critical business insights can be discovered by using this calculation? What is the applied benefit of the topic? How would a professional use it to succeed?
- As we work on the related activities, remember… [give them a tip, or encourage them to help each other, or contact you during office hours if they need help.]
- Some kind of closing: I look forward to reviewing your work and have a great week.

Course Intro Notes

- Welcome to Time Series. My name is *Son Nguyen* and I will be your professor for the next ten weeks.

- Why/Motivation:

  - Forecasting is at the heart of actuarial science. Actuaries study datasets of claims, expense, investment, and other economic processes developing over time in order to build models for insurance systems that will operate in the future.

  - These datasets are called time series.

- What to cover:

  - In this course, we will

    (1) Start with some basic concepts of time series such as white noise, stationary and autocorrelation

    (2) We then study some of the most popular Time series Models such as Random Walks, ARMA, ARIMA and how to use these models for forecasting

(3) We will also learn how to use R to help us implement all of these models and Analysis. R is the most popular statistical language and also the language used in the SOA and the CAS.

- I began my work in this area as a *PhD student* at *Ohio University more than 10 years ago*. Since then, I have *been teaching and doing research in the area of statistical learning and data mining.* I joined Bryant University in 2016, where I am a member of the department teaching courses in Applied Data Mining, Statistical Analysis with R.

- Thank you for joining me for this course. I look forward to working with you online.

Week 1 Notes: *Basic concepts in Time Series*

- *This week, we will basic concepts in Time Series including concepts of White Noise, Stationary, Autocorrrelation. These concepts will build a foundation for our time series analysis and modeling in the coming weeks.*

- This week we also introduce R and Rstudio and discuss how we can use this platform help us understanding and implementing time series techniques.

- In the pod assignment, we will set up R and Rstudio and write our first lines of R codes to explore a time series.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 2 Notes: Smoothing Techniques.

- Smoothing is usually done to help us better see patterns, trends for example, in time series. Generally smooth out the irregular roughness to see a clearer signal.

- This week we will study two of the most popular smoothing techniques: Moving Average Smoothing and Exponential Smoothing. We will discuss the pros and cons of these two techniques and also how to apply these techniques in real datasets.

- In the pod assignments, we will us R to implement the

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 3 Notes: Exponential Smoothing Models

- *Exponential smoothing was proposed in the late 1950s, and has motivated some of the most successful forecasting methods. This week we will discuss several exponential smoothing models and how to use these models for forecasting.*

- *We will see that exponential smoothing methods can be very flexible and capable of handling time series with trend, seasonality or both.*

- In the pod assignments, we will implement exponential smoothing in both real-life datasets and our own simulated datasets.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 4 Notes: Time series regression models

- Linear Regression has been a powerful model in predictive modeling and is still one of the most popular model used nowaday.

- This week we discuss how linear regression models can be applied on time series. You will see that the basic concept is that we forecast the time series of interest assuming that it has a linear relationship with other time series.

- We will also compare some fundamental different between a regular regression model and a time series regression model.

- As usual, in the pod assignment, we will bring in a real-life datasets to perform the regression model.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 5 Notes Random Walks Models

- This week we will look at one of the simplest and yet most important models in time series forecasting: the random walk model.

- The random walk model simply says that the difference between any consecutive values in a time series is always a random quantity come from the same

distribution.

- How such a simple model can help us understand a complex system and have such an immense impact and applications in many fields including engineering, computer science, biology and social sciences? This week, we will discuss the important characteristics of this model and implement it in real-life dataset.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 6 Notes Autoregressive models - AR

- This week, we will study one of the most widely used model in time series forecasting: the autoregressive model.

- If in linear regression, we forecast the variable of interest using a linear combination of predictors. In an autoregression model, we forecast the variable of interest using a linear combination of past values of the variable.

- The term autoregression indicates that it is a regression of the variable against itself .

- In the implementation of the model, you will see that autoregressive models are remarkably flexible at handling a wide range of different time series patterns.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 7 Notes: ARIMA Models (1)

Week 8 Notes: ARIMA Models (2)

- We are now ready to talk about one of the most important and most widely - used model in time series: the ARIMA model.

- This model can also be considered a family of models as it include: the random walk model, the auregressive model and the moving average models. Even a white-noise is a special case of the ARIMA model.

- In the pod assignments, we will demonstrate that this model is extremely flexible and can handle a variety of time series data.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 9 Notes: Machine Learning Models (1)

Week 10: Machine Learning Models (2)

- With the availability of data and computational powers, machine Learning Models have become so popular in the field of predictive modeling in the past decades.

- In this section, we will see how machine learning models can be used for time series forecasting and what are the key difference between applying machine learning models on a time series dataset and on a cross-sectional dataset.

- We then implement several machine learning models in real-life time series datasets in R.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

—

- What to cover:

    - In this course, we will cover three components. We will study supervised learning techniques where we want to predict a future value of an interested variables.
    - We also investigate unsupervised learning techniques which helps us understand the structure of data.
    - We learn how to use R to help us implement all of these techniques. R is the most popular statistical language and also the language used in the SOA predictive analytics exam

- The goal of this course is to equip you will all the needed tools to build top quality risk models and help you prepare for the SRM and PA exam of the SOA and the MAS-II exam of the CAS.

- I began my work in this area as a *PhD student* at *Ohio University more than 10 years ago*. Since then, I have *been teaching and doing research in the area of statistical learning and data mining.* I joined Bryant University in 2016, where I am a member of the department teaching courses in Applied Data Mining, Statistical Analysis with R.

- Data is everywhere. The market size for skills in Statistics and Data Analytics is expected to grow more than 200% from 2021 by 2026. The skills to work with

data is needed more than ever.

- Thank you for joining me for this course. I look forward to working with you online.

Week 1 Notes

- *This week, we will look at the simplest yet effective model of them all*: Simple Linear Model.

- Linear Models have a long history and still one of the most popular predictive model nowaday.

- Linear models is attractive in many ways: it has an elegant solution, it is very flexible to apply and it is so easy to interpret.

- Understanding linear models will help us better understand many other advanged supervised learning models.

- Our reading assignments and the class notes will walk you through the theory of simple linear models and also give you a guideline to solve the problelms in the assignments.

- In the pod assignment, you will have an opportunity to implement linear models in R with a real-life example.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 2 Notes

- *This week, we will look at a natural extension of linear model*: Multiple Linear Model.

- If in simple linear model, we make a prediction using only one predictors, in multiple linear models, we are allowed to have multiple predictors. This makes multiple linear model is so much more versatile than SLM.

- This may be the model you try first as a risk modeller. And iIf there is one model you have to understand fully, I would recommend multiple linear models.

- We will discuss different aspect of building, evaluate and interpret the results of MLM. In the pod assignment, we will strengthen our understanding of this model by solving real-life examples using MLM with the help of R.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 3 Notes

- *MLM has some limitation: it cant work with a binary response, for example, predicting if a transaction is a fraud. It cant also capture non-linear relationships between predictors and response.*

- *This week, we will discuss a class of models that overcomes these limitations: generalized linear models.*

- GLM is a very powerful model that give the modeller the freedom to choose a relationship between between predictors and the response. GLM is a class of models including many famous model such as MLM, logistic regression, and Poisson regression.

- We will discuss the theory behind GLM and how to choose the right GLM model for a given dataset. We will study two important GLM models: Logistic Regression and Poission Regression. In the pod assignment, we will use R to build GLM models with real-life dataset.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 4 Notes

- This week we will look at special type of data: time series. A time series is a sequence of data indexed in time order. You will find time series data everywhere, and especially in Financial and Economic applications. For example: a list of apple stock prices in the last 30 days is a time series.

- Business can use time series to forecast future demand to plan inventory and optimize operations.

  Government can use times series to understand and monitor the performance of key economic indicators such as GDP and inflation rates.

- We will cover some fundamental concepts of time series analysis such as time series decomposition and autocorrelations. We will also discuss forecasting techniques to help you build an effective model for future prediction.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 5 Notes

- This week we will study the most visual and intuitive model of them all: decision trees.

- Why is this topic important? Decision Trees is one of a few models that we can actually visualize. This makes decision trees extremely easy to interpret and to communicate their results.

- We will talk about how to growth a decision tree starting from the very first branch. In the pod assignments, you will train and evaluate decision trees on a real-life dataset.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 6 Notes

- This week we will discuss one of the most important ideas in predictive modeling: ensemble learning. We will see how we can form a committee of models to make the final improved prediction. In particular, we will form a forest from multiple decision trees.

- Ensemble methods have been proven to be one of the best way to build an accurate model as we often see in real life that a team usually outperform an individual. The winning models in modeling competitions are usually in the form of an ensemble model.

- You will learn how a forest are formed and its advantages comparing to decision trees. In the pod assignments, you will train and evaluate a random forest on a real-life dataset.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 7 Notes: KNN

- “Tell me who your friends are, and I will tell you who you are”. This is another life wisdom that turns into an effective predictive model: K-Nearest Neighbors or you can call it K-best friends.

- KNN is very simple and easy to understand. It has proven to be very effective in many different domains.

- You will learn different ways to measure the closeness between neighbors/friends and how to compute a

prediction with KNN. In the pod assignments, you will train and evaluate a KNN model on a real-life dataset.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 8 Notes: Clustering

- This week we will learn one of the most important method in unsupervised learning: clustering.

- Clustering is grouping data points into categories. Clustering has many applications. Insurance company can use clustering to divide their potential clients into groups and then design suitable insurance products for each group.

- There are plenty of clustering methods. We will look at two of the most popular ones: Hierarchical clustering and K-mearns clustering. You will learn how these methods work and implement them in R with real-life datasets.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

- I look forward to reviewing your work and have a great week.

Week 9 Notes: PCA

- This week we will look at one of the most popular methods in machine learning: Principal Component Analysis.

- PCA turns the a dataset of thousand variables to a few variables without losing important information. This comes in extremely handy with all the larger datasets we have nowadays.

- It is a very common practice to use PCA before applying other machine learning methods such clustering, modeling or data visualization.

- We will discuss the fundamental idea behind PCA and see its applications with real-life examples.

- Make sure you reach out to your pod mates for the Pod assignment and please see me during office hours if you need any help.

Week 10: Final

- Welcome to the last week of the class. This week we will review and strengthen our understanding of all topics we have covered.

- We will practice problems that likely appear on actuarial exams and prepare you for the final exam.

- We will hold live sessions to go over all the questions that you may have. Make sure you tune in for the session and will see you there.