Proposal for the Saudi Digital Academy
Prepared by Coding Dojo
March 2020.

Data Science Online Immersive Bootcamp

Objective of the Program:
Walk away with a work applicable skill set to apply the Data Science process methodologies and tools to solve real-world problems in business and academia. Students will be able to apply data science techniques and approach inside of real-world organizations. Starting with first principles, students should finish with a skill set which makes them competitive in the workforce as a junior data scientist or data analyst.

- Entry level students with STEM knowledge gain a comprehensive understanding of the fundamentals of Data Science

Expected outcomes:

- Understand the Data Science Process and how to apply it in various situations
- Develop Data Science solutions using SQL, Tableau, R, and Python
- Build predictive models by leveraging statistical and machine learning
- Understand how business problems can be solved using Data Science techniques

Commitment

- 14 weeks on-site: Instructors and Students on-site 40hrs/wk for 14 weeks
- Mixed format of lectures, videos, reading, assignments and projects
- Full time, 40 hours/week

Landing Page

- Link: https://www.codingdojo.com/data-science-bootcamp

Admissions Process Prerequisites

- A desire to learn data science through top-quality instruction
- A strong interest in solving complex problems through data analysis and data science in a systematic, principled way
- STEM experience a plus
- Familiarity w/ linear Algebra and basic Python and R programming

Requirements to Join:

- Personal Laptop
- Stable Internet Connection
- Modern web browser such as Firefox or Chrome
- Complete prep-work before day-1 of class

Enrollment Process

1. Apply, fill-out enrollment form
2. Skills Assessment
3. Interview
4. Submit tuition in Learning Platform to finalize enrollment
5. Receive prep-work via email

Attendance Policy

- 80% attendance of classes, arriving by 9am for group activity
- Weekly collaboration sessions required

Graduation Requirements

- Greater than 80% attendance throughout the entire course.
- Complete 70% or more of assigned non-optional pass / fail assignments or projects
- Passing of the Pass / Fail assignments or projects requires solving or producing the simulations assigned. Additional feedback will be provided by the instructor to each student for each assignment / project about the areas where they met, exceeded or did not meet the requirements of the assignment / project.

Course Title:

- Data Science Immersive Bootcamp
- 14- weeks; Instructor and Students on-site 40hrs/wk, hybrid instruction incorporating Learning Management System platform
- **Lecture Hours:** 10 hours per week
- **Lab Hours:** 30 hours per week
- **Total Course Hours:** 560 hours

Curriculum Modules

Week 1 **Intro to Data Science and Business Intelligence**	**Introduction to using Data** **Learn about data science and the data science process. Discover various data types and formats. Learn about spreadsheet basics and how to interact with and build them. Build basic business intelligence applications MS Excel.** Monday: Introduction to Data Lesson Objective 1: Understand the history of data Lesson Objective 2: Understand the history of data science Lesson Objective 3: Understand what data science, data analysis, software development and data engineering are and how they relate Assignment: Reading Tuesday: Introduction to Data Analysis With Spreadsheets Lesson Objective 1: Understand the spreadsheet concept Lesson Objective 2: Import and Manipulate data in MS Excel Lesson Objective 3: Build basic visualizations in MS Excel Assignment: Intro to Stock Market Data

	Wednesday: Data Analysis With Spreadsheets Lesson Objective 1: Able to format an Excel document Lesson Objective 2: Build basic functions and formulas Lesson Objective 3: Build pivot tables to explore data Assignment: First Stock Portfolio Thursday: Data Analysis with Spreadsheets Lesson Objective 1: Build logic and lookup functions Lesson Objective 2: Use and understand Excel financial functions Lesson Objective 3: Understand sensitivity analysis and the goal seek command Assignment: Advanced Stock Portfolio Friday: Assignment: Build an Excel dashboard to present financial information about a startup idea **Technologies:** MS Excel
Week 2 **Business Intelligence with Databases**	**Working with Databases** **Learn how to interact with the most common data storage systems using the structured query language SQL.** Monday: Lesson Objective 1: Understand why databases are useful Lesson Objective 2: Understand how SQL is used to interact with databases Lesson Objective 3: Able to build basic components of databases Assignment: First DB Assignment Tuesday: Lesson Objective 1: Able to build basic analytic queries Lesson Objective 2: Able to construct SQL queries from a problem statement Lesson Objective 3: Execute SQL queries from an existing database client Assignment: SQL External Resources Wednesday: Lesson Objective 1: Able to build SQL queries that involve select Lesson Objective 2: Able to build SQL queries that involve from Lesson Objective 3: Able to build SQL queries that involve where, limit, sorting Assignment: First Query and Aggregation assignment Thursday: Lesson Objective 1: Build SQL queries the involve table joins Lesson Objective 2: Build SQL queries that require functions Lesson Objective 3: Build SQL queries that require subfunctions Assignment: Advanced SQL External resources

	Friday: Assignment: Creating and Populating a Database **Technologies:** SQL, Superset
Week 3 **Business Intelligence**	**Displaying and Presenting Data Using Business Intelligence Tools** **Use Tableau and Superset to present data to decision makers** Monday: Lesson Objective 1: Understand what business intelligence is and how it can improve business processes Lesson Objective 2: Identify and describe business intelligence tools Lesson Objective 3: Understand how business intelligence tools fit into a typical data ecosystem Assignment: Connecting your Database to a Dashboard Tuesday: Lesson Objective 1: Understand how databases connect to business intelligence systems Lesson Objective 2: Able to build basic plots in superset from a connected database Lesson Objective 3: Publish a basic analytic from superset Assignment: Advanced Data Analysis Dashboards Wednesday: Lesson Objective 1: Understand what Tableau is and how to create an account on Tableau Public Lesson Objective 2: Able to build basic plots form an imported text file Lesson Objective 3: Import and connect datasets to Tableau Assignment: First Tableau dashboard Thursday: Lesson Objective 1: Build custom visualizations using the dashboard Lesson Objective 2: Create a story using Tableau Lesson Objective 3: Publish a story to the internet using Tableau Public Assignment: First Tableau public story Friday: Assignment: Advanced Tableau story on economics data **Technologies:** Superset and Tableau
Week 4 **Data Munging, Analysis, and**	**Foundations of Data Munging** **Using the R statistical programming language to clean, prepare, and visualize data.**

Visualization using R	Monday: Introduction to Scripting Languages through R Lesson Objective 1: Understanding R and the Rstudio Interface Lesson Objective 2: Load and use packages Lesson Objective 3: Create custom R objects, functions and scripts Assignment: Getting Comfortable with R Tuesday: Programming in R Lesson Objective 1: Assembling and Disassembling Data sets Lesson Objective 2: Loading and unloading data sets Lesson Objective 3: Modifying Data in an existing object Assignment: Manipulating Data Wednesday: Programming in R Lesson Objective 1: Understanding If statements and loops Lesson Objective 2: Writing vectorized code that is fast and flexible Lesson Objective 3: Debugging your code first principles Assignment: Vectorized Code Thursday: Simulation and Plotting Lesson Objective 1: Understanding Simulations Lesson Objective 2: Building custom graphs using plot function Lesson Objective 3: Storytelling with visualizations Assignment: Simulations Friday: Assignment: Build a simulation of a casino game and provide custom plots to show profit over time **Technologies:** R and Ggplot2
Week 5 **Advanced data preparation using the tidyverse**	**Advanced-Data Preparation** **Using the tidyverse collection of packages to more efficiently analyze data.** Monday: Lesson Objective 1: Understanding the Tidyverse Lesson Objective 2: What is Munging data and why is it important Lesson Objective 3: Build custom visualizations using ggplot Assignment: Tidyverse plotting Tuesday: Lesson Objective 1: Transform a data set to reflect a business need using Tidyverse functions Lesson Objective 2: Understanding Exploratory Data Analysis Lesson Objective 3: Performing EDA on an unknown problem/dataset Assignment: Performing my first EDA

	Wednesday: Lesson Objective 1: Understanding Tidy Data Lesson Objective 2: Manipulating unstructured data and making it tidy Lesson Objective 3: Tidying data from a relational database Assignment: Data Cleaning 101 Thursday: Lesson Objective 1: Tidying string data Lesson Objective 2: Tidying date time data Lesson Objective 3: Tidying factor data Assignment: Data Cleaning 201 Friday: Assignment: First full munge project **Technologies:** R and Tidyverse
Week 6 **Statistical Analysis using R**	**Statistical Analysis (Modeling)** **Learn summary statistical and probability concepts. Learn how to use linear regression to explore relationships.** Monday: Lesson Objective 1: Understanding the different types of models Lesson Objective 2: Understanding the bias/variance tradeoff Lesson Objective 3: Understanding training, validation and testing data sets Assignment: Modeling Basics Tuesday: Lesson Objective 1: Understanding basic probability concepts Lesson Objective 2: Understanding probabilistic models Lesson Objective 3: Understanding the basics of decision analysis Assignment: Probability decisions and first stochastic models Wednesday: Lesson Objective 1: Understanding linear optimization Lesson Objective 2: Understanding Markov Processes Lesson Objective 3: Building a basic model Assignment: First MDP and LP Thursday: Lesson Objective 1: Building custom reactive shiny apps Lesson Objective 2: Building Custom documents with Rmarkdown for presentation Lesson Objective 3: building an API using plummer Assignment; Turning models into results

	Friday Assignment: Full R project with deterministic and Probabilistic models **Technologies:** R and Tidyverse
Week 7 **Scientific Programming using Python**	**Using Python for Scientific Applications** **Learn the basics of the Python programming language. Learn how to use Python for scientific and mathematical applications. Use python for basic linear algebra.** Monday: Lesson Objective 1: Understand how Python can be used for scientific analysis Lesson Objective 2: Able to use Python to solve basic mathematical problems Lesson Objective 3: Able to use Python to solve basic logic problems Assignment: Math Refresher Tuesday: Lesson Objective 1: Understand how Numpy is useful for solving mathematical problems Lesson Objective 2: Able to solve specific linear algebra problems with Numpy Lesson Objective 3: Able to solve a linear systems of equations using Numpy Assignment: Numpy 101 Wednesday: Lesson Objective 1: Understand how random numbers are generated using scipy Lesson Objective 2: Able to use random numbers to solve logic problems Lesson Objective 3: Build function that use random number generators Assignment: Random Variable creation Thursday: Lesson Objective 1: Able to build OOP method and classes for scientific applications Lesson Objective 2: Able to create an application based package for scientific applications Lesson Objective 3: Understand how to incorporate applications into a problem setting Assignment: Creating Packages Friday: Assignment: Python for Data Science 101 **Technologies:** Python Numpy Scipy

Week 8 **Data Munging, Analysis, and Visualization using Python**	**Python for Data Analysis and Cleaning** **Learn to use Python for Data Science related tasks.** Monday: Lesson Objective 1: Understand how to ingest data into Python Lesson Objective 2: Able to identify and ingest different data types Lesson Objective 3: Able to extract data from an API Assignment: Using Pandas Tuesday: Lesson Objective 1: Able to perform basic manipulation of data Lesson Objective 2: Extract basic insight from a dataset Lesson Objective 3: Able to calculate summary statistics Assignment: Cleaning Data in Python Wednesday: Lesson Objective 1: Able to build insightful visualizations Lesson Objective 2: Understand that different types of viz for different applications Lesson Objective 3: Able to use matplotlib for basic visualizations Assignment: Basic data Manipulation Thursday: Lesson Objective 1: Able to use seaborn for efficient visualization Lesson Objective 2: Able to use a balance of summary statistics and visualization to extract insight. Lesson Objective 3: Able to communicate insight that is uncovered from data Assignment: Advanced Visualization Friday: Assignment: First Python Project **Technologies:** Python Pandas Matplotlib Seaborn
Week 9 **Exploratory Data Analysis using Python**	**Comprehensive Exploratory Data Analysis** **Use Python to explore and analyze various datasets. Build and deploy a dash app.** Monday: Lesson Objective 1: Understand the impacts of unclean data Lesson Objective 2: Able to replace missing values Lesson Objective 3: Able to replace incomplete values Assignment: Cleaning Data in Python

	Tuesday: Lesson Objective 1: Understand the statistical implication of replacing missing values Lesson Objective 2: Calculate summary statistics on selected data Lesson Objective 3: Able to visualize summary statistics of selected data Assignment: Basic Statistics HW Wednesday: Lesson Objective 1: Understand how to build interactive components using Dash Lesson Objective 2: Able to build a basic dash application Lesson Objective 3: Able to deploy a basic dash application Assignment: Building my first Dash App Thursday: Lesson Objective 1: Able to build a user facing dash app Lesson Objective 2: Able to integrate EDA insights into a dash app Lesson Objective 3: Able to integrate summary statistics into a dash app Assignment: Advanced Dash App Friday: Assignment: Python Munging, Cleaning and Presenting Data **Technologies:** Python Dash
Week 10 **Introduction to Machine Learning**	**Comprehensive Exploratory Data Analysis** **Learn the basics of Machine Learning Concepts. Learn about a common set of Machine Learning Techniques for classification and regression.** Monday: Lesson Objective 1: Understand what Machine Learning is Lesson Objective 2: Understand the types of Machine Learning Lesson Objective 3: Understand the difference between Machine Learning and Data Science Assignment: ML Basics Tuesday: Lesson Objective 1: Understand what a heuristic model is Lesson Objective 2: Able to build a heuristic model Lesson Objective 3: Able to evaluate the quality of a heuristic model Assignment: Creating my first Heuristic Model Wednesday: Lesson Objective 1: Understand what a cost function is

	Lesson Objective 2: Able to build a cost function for linear regression Lesson Objective 3: Understand what gradient descent/nonlinear optimization are Assignment: Cost Function Implementation Thursday: Lesson Objective 1: Able to build a linear regression model Lesson Objective 2: Able to apply gradient descent to sum of squared error Lesson Objective 3: Able to predict an outcome using a linear model Assignment: My first ML model Friday: Assignment: Predictive and Inference Model **Technologies:** Python & R Scikit-learn - Python Parsnip - R
Week 11 **Intermediate Machine Learning**	**Classification Problems in Supervised Machine Learning** **Learn about more advanced machine learning techniques of ensemble Methods.** Monday: Lesson Objective 1: Understand the Scikit-learn API Lesson Objective 2: Able to build a regression model with Scikit learn Lesson Objective 3: Able to predict an outcome with Scikit learn Assignment: Python Prediction 101 Tuesday: Lesson Objective 1: Understand the Parsnip API Lesson Objective 2: Able to build a regression model with Parsnip Lesson Objective 3: Able to predict an outcome with Parsnip Assignment: R Prediction 101 Wednesday: Lesson Objective 1: Understand what a classification ML problem is Lesson Objective 2: Understand how to calculate the quality of a classification model Lesson Objective 3: Able to calculate the precision and recall of a classification model Assignment: Classification Model Assignment Thursday: Lesson Objective 1: Understand how a logistic regression is used for classification Lesson Objective 2: Able to build a logistic regression using scikit-learn

	Lesson Objective 3: Able to build a logistic regression using parsnip Assignment: My first Logit Model Friday: Assignment: Full blown classification problem **Technologies:** Python & R Scikit-learn - Python Parsnip - R
Week 12 **Intermediate Machine Learning**	**Special Topics in Machine learning** **Learn about more advanced machine learning techniques of ensemble Methods.** Monday: Lesson Objective 1: Understand Feature Engineering Lesson Objective 2: Able to build custom features to improve a regression model Lesson Objective 3: Able to build custom features to improve a classification model Assignment: Feature creation and integration Tuesday: Lesson Objective 1: Understand Naive Bayes Classification Lesson Objective 2: Able to build a naive bayes using scikit-learn Lesson Objective 3: Able to predict an outcome using naive bayes using scikit-learn Assignment: Naive Bayes model Wednesday: Lesson Objective 1: Understand ensemble models are Lesson Objective 2: Understand boosting and bagging Lesson Objective 3: Understand CART models and how they are used for bagging Assignment: Building my first Ensemble model Thursday: Lesson Objective 1: Understand how a decision tree works for classification Lesson Objective 2: Understand how a random forest works for classification Lesson Objective 3: Able to build a random forest for classification using scikit-learn and parsnip Assignment: Decision Trees and Random Forest models deployed Friday: Assignment: Build a classification Ensemble model

	Technologies: Python & R Scikit-learn - Python Parsnip - R
Week 13 **Basic Data Engineering**	**Comprehensive Exploratory Data Analysis** **Learn about building and deploying data pipelines and models.** Monday: Lesson Objective 1: Understand what data pipeline is Lesson Objective 2: Understand where to use a data pipeline Lesson Objective 3: Able to build a database for data pipelines Assignment: Building and creating a pipeline Tuesday: Lesson Objective 1: Able to build a basic data pipeline with Python Lesson Objective 2: Able to deploy a basic data pipeline with flask Lesson Objective 3: Able to use a basic data pipeline with Python Assignment: My first Flask app Wednesday: Lesson Objective 1: Able to build a basic data pipeline with R Lesson Objective 2: Able to deploy a basic data pipeline with plumber Lesson Objective 3: Able to use a basic data pipeline with R Assignment: My first Plummer App Thursday: Lesson Objective 1: Understand how nested pipelines work Lesson Objective 2: Able to build conditional data pipelines Lesson Objective 3: Able to build complex data pipeline interactions Assignment: Bridging different pipelines Friday: Assignment: Building a complex pipeline **Technologies:** Python & R Flask - Python Plumber - R
Week 14 **Project Week**	**Comprehensive Data Science Project** Project: Build a comprehensive Data Science Project from scratch. **Technologies:** Python & R

Online Learning Plan

Due to the unforeseen current health challenges that we are facing, we will deliver the programming bootcamp through online services. Listed below are the specific provisions we have made to provide the best online experience for your candidates.

1. **All lectures will be streamed live during the 0900-1600 AST workday and will be coordinated on a weekly and daily basis with the instructors and teaching assistants.**
2. **Zoom Pro will be used for, lectures, webinars and 1-on-1 sessions**
3. **Walkabout https://www.walkaboutco.com/ We will be implementing a custom Walkabout virtual workplace for the bootcamp to provide presence and a place to organize and collaborate with students.**
4. **Shift Instructors: The course will have 2 instructors dedicated to the course and 2 TAs to provide personal and curated service and learning in the local time zone (AST).**
5. **Recorded Lectures will be cataloged and shared with all students so that can rewatch lectures for up to 2 years.**
6. **Access to all online course material, our online team is developing new lectures and videos daily. SDA students will have access to all of the materials for 2 years.**
7. **Daily activities will be held by TAs and Instructors in addition to the lectures, and will also be recorded.**

Instructional Staff:

Project Lead and Senior Instructor: Daniel H. Oostra
https://www.linkedin.com/in/danieloostra/

Lead Instructor: Willem Pretorius
https://www.linkedin.com/in/willem-pretorius

Two (2) Teaching Assistants
TBD

Client Management and Executive Support

Kiana Pan, VP Partnerships and Customer Success
https://www.linkedin.com/in/kiana-pan-b68b2282/

Cost Estimate for Data Science Bootcamp:
Pre-course for 30 applicants: $3,750 USD.

For the delivery of this bootcamp, Coding Dojo will charge $10,000 per student.
Total cost for a 25 person cohort: $250,000 USD.

Additional related costs:
Walkabout Online Services: $900.00

Total Project Budget: $254,650 USD