شـهادة إتـمـام تـدريب
Certificate of Completion

تبارك الأكاديمية السعودية الرقمية لــــ

Saudi Digital Academy Wishes to Congratulate

مرام الشهري

Maram Alshehri

إتمام معسكر همة لعلم البيانات بنجاح

On Successfully Completing Data Science BootCamp

بعد رحلة استمرت لمدة 14 أسبوعًا

In A Journey That Lasted 14 weeks

في الفترة من 15 شعبان 1442

From 28 March 2021

إلى 5 ذو الحجة 1442

To 15 July 2021

تم إصدارها بتاريخ:

Issue Date:

5 ذو الحجة 1442

15 July 2021

مع تمنياتنا بدوام التوفيق والنجاح

Wishing You A Continuous Success
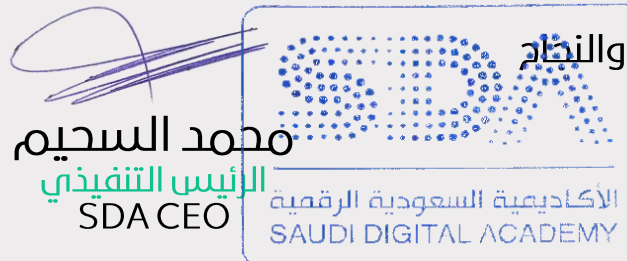
محمد السحيم
الرئيس التنفيذي
SDA CEO

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

رقم الشهادة:202107114

BootCamp Plan

Introduction to Data Science and Business Intelligence	**Week 1** • History of data. • Understand the relationship between data science, data analysis, and data engineering. • MS Excel: Data manipulation and visualization, Basic and advanced functions includes logic and lookup, Pivot tables, Sensitivity analysis.	**Week 2** • SQL for Data Analysis. • Databases objectives and application • SQL Analytics queries includes functions, subfunctions, and table joins. • Relational databases and non relational databases.	**Week 3** • NoSQL Databases: Comparison between SQL and NoSQL, MongoDB Compass and MongoDB pipeline • Business Intelligence: Overview on different BI tools and how they fit into data ecosystem. • Tableau: Import and connect datasets, build custom visualizations using the dashboard, Tableau stories, publish using Tableau public.
The Statistical Programming Language R	**Week 4** • Intro to scripting language: R language and Rstudio interface. Load and use packages, Functions and custom objects. • Assembling and Disassembling Data sets, data modification. • Conditional statements and data vectorization in R. • Simulation and storytelling with visualization.	**Week 5** • Intro to Tidyverse, data mugging with Tideyverse. • Custom data visualization using ggplot. • Understanding Exploratory Data Analysis (EDA) and performing EDA on an unknown problem/dataset. • Understanding Tidy Data, manipulating unstructured data making it tidy, tidy data in relational databases. • Tidying different data types such as string, date time, and factor data.	**Week 6** • Intro to modeling and basic concepts: different types of models, bias/variance tradeoff, understanding of training, validation and testing sets • Probability and decision analysis: basic probability and statistical concepts, probabilistic models. • Optimization: Linear optimization and Markov processes. • Sharing and presenting models: Reactive shiny apps, custom documents with Rmarkdown, and APIs.
Introduction to Python	**Week 7** • Intro to programing using anaconda and Jupyter notebook. • Intro to NumPy, and linear equations. • Loops, conditional statements, and user-defined functions on Python. • Overview on Python packages.	**Week 8** • Ingest data into Python and extract data from an API. • Data manipulation and wrangling. Summary statistical analysis. • Data visualization using Matplotlib and Seaborn. • EDA project with insightful visualization.	**Week 9** • Understand the impacts of unclean data. • Deal with missing and incomplete values with an understanding of the statistical implication of replacing the missing values. • Scientific visualization using dash application and integrate EDA insights into a dash app.
Introduction to Machine Learning	**Week 10** • Intro to machine learning. • Build a heuristic model and able to evaluate the quality of their work. • Understand the cost function, gradient and decent/nonlinear optimization. • Build linear regression model and evaluate the prediction using deferent evaluation metrics.	**Week 11** • Understand the Scikit-learn API and its applications. • Understand the quantitative models. • Build logistic regression for classification using Scikit-learn. • Apply feature engineering and able to improve regression and classification models.	**Week 12** • Understand the CART models. • Overview on ensemble modeling and understanding of boosting and bagging techniques. • Understand the Principal Component Analysis (PCA) and able to build PCA using scikit-learn. • Understand how decision tree and random forest work for classification. • Clustering technique such as K-Means.
	Week 13 • Overview on basic data engineering techniques. • Understand data pipeline and deploy it with Python , R, Flask, and Plumber. • Build conditional, nested and complex data pipeline interactions.		
Week 14	**Capstone Project**		