

DATA SCIENCE FOUNDATION BOOTCAMP

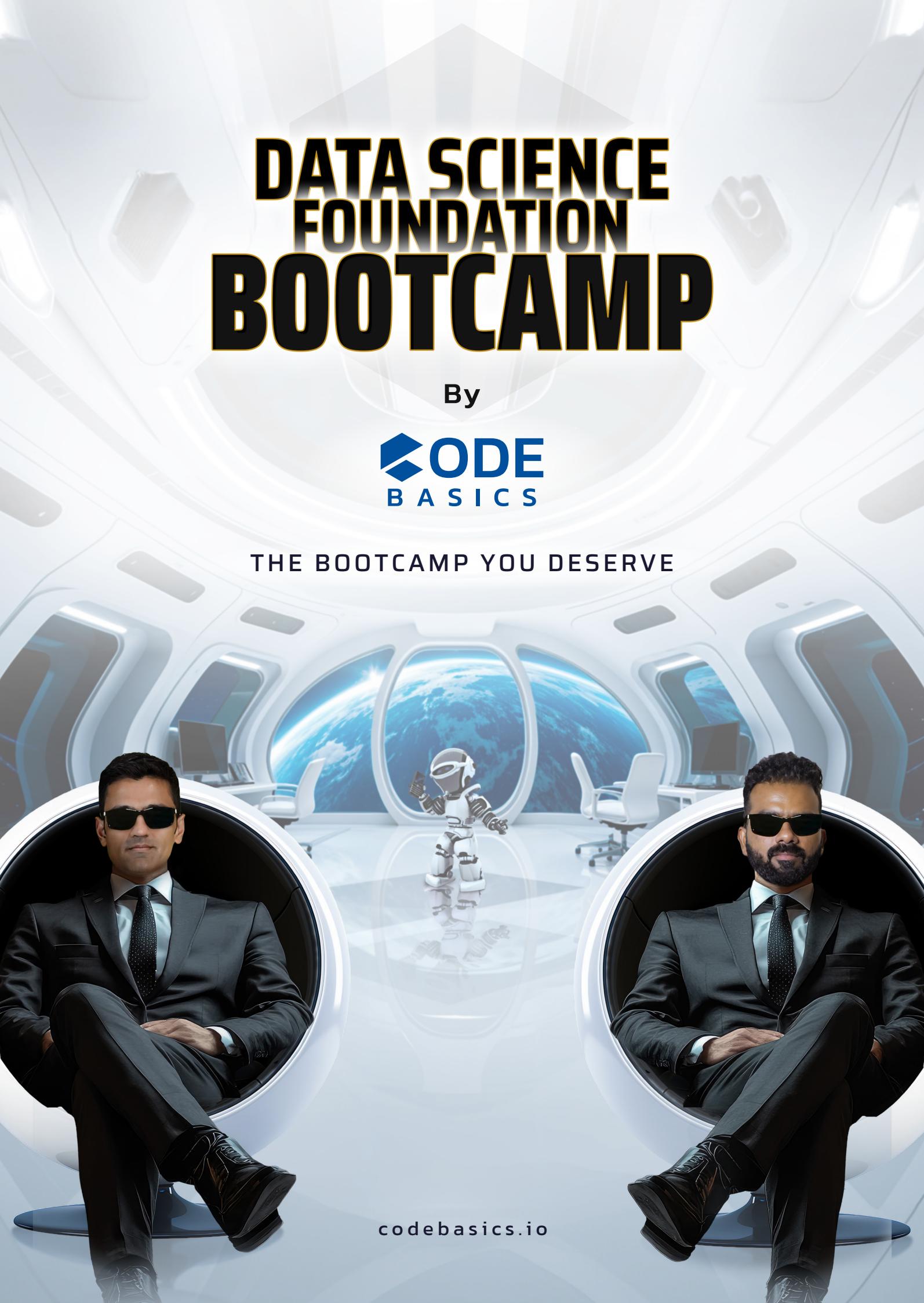
By



THE BOOTCAMP YOU DESERVE



codebasics.io



CODEBASICS STORY



In 2016, **Dhaval Patel**, the founder of codebasics, created the '**codebasics**' YouTube channel to share his knowledge as a means to gain mental peace during a significant health struggle. He discovered his passion for teaching and realised his unique ability to simplify complex topics. The love and blessings from thousands of students started to heal his health issues. This transformative experience soon became his life's purpose and even led him to consider leaving his lucrative job at Bloomberg, New York.

In 2020, one of Dhaval's subscribers, **Hemanand Vadivel, a data analytics expert** working in Europe, wished to contribute to the codebasics community as a 'payback' for what he learned from the codebasics channel. In 2021, Hemanand began creating impactful videos for codebasics, and it helped him heal from the trauma of losing family members to Covid-19. Seeing their shared values, Dhaval proposed a collaboration: creating the best, most affordable Power BI course on the planet.



JUNE 2022

After six months of dedicated collaboration, they released the Power BI course in June 2022, which was met with overwhelming positivity. By this time, both Dhaval and Hemanand had already left their high-paying jobs to create 'codebasics' as a full-time learning tech company. They made this decision after discovering their 'Ikigai', their true purpose: to make a meaningful impact through education, without any gimmicks.

In the era of commoditized education, where millions of courses are offered for profit but lack true value, Dhaval and Hemanand prioritized providing authentic, job-relevant learning experiences with honest pricing. Besides job-related affordable courses, codebasics offers high-quality free content and conducts several initiatives to support aspirants including mothers returning to work. This embodies their core vision of providing authentic job-related learning experiences to everyone, irrespective of their financial condition.

MAY 2024

Codebasics has more than
1 Million+ learners
in their community and strongly marching
forward with a vision to unleash at least
5 Million+ potentials by 2025.

YOUR TEACHERS ARE INDUSTRY EXPERTS



Dhaval Patel

**Data Entrepreneur (12+ Years),
YouTuber,
Ex - Bloomberg, NVIDIA**

I have 17 years of experience in Programming and Data Science working for big tech companies like NVIDIA and Bloomberg. I also run a famous YouTube channel called Codebasics where I pursue my passion for teaching.



Hemanand Vadivel

**Ex- Data Analytics Manager,
8+ Years in Europe, Microsoft Certified,
Certified Supply Chain Professional**

I'm a Mechanical Engineer who transitioned to a full-time Data & Analytics Manager in the UK & Germany. I have delivered 30+ analytics projects over 15+ countries and trained professionals at different levels to equip them with valuable analytics skills.



HOW WE TEACH



Learn

Learn from the industry experts and the YouTuber, Dhaval Patel who is loved and known for his simple explanations.

With simple explanations, project-based learning, and real-world business context, we help learners enjoy a comprehensive learning experience that stays with them for a lifetime.

Engage



Our courses have a higher completion rate than average for a reason.

With cinematic storytelling, immersive visual effects, and relatable examples, we ensure that learning remains effective, impactful, and engaging!



Get Challenged

Learning technology is like learning swimming - you can truly learn it only when you practice. That's why we offer a range of challenging quizzes and exercises that help learners apply what they have learned.

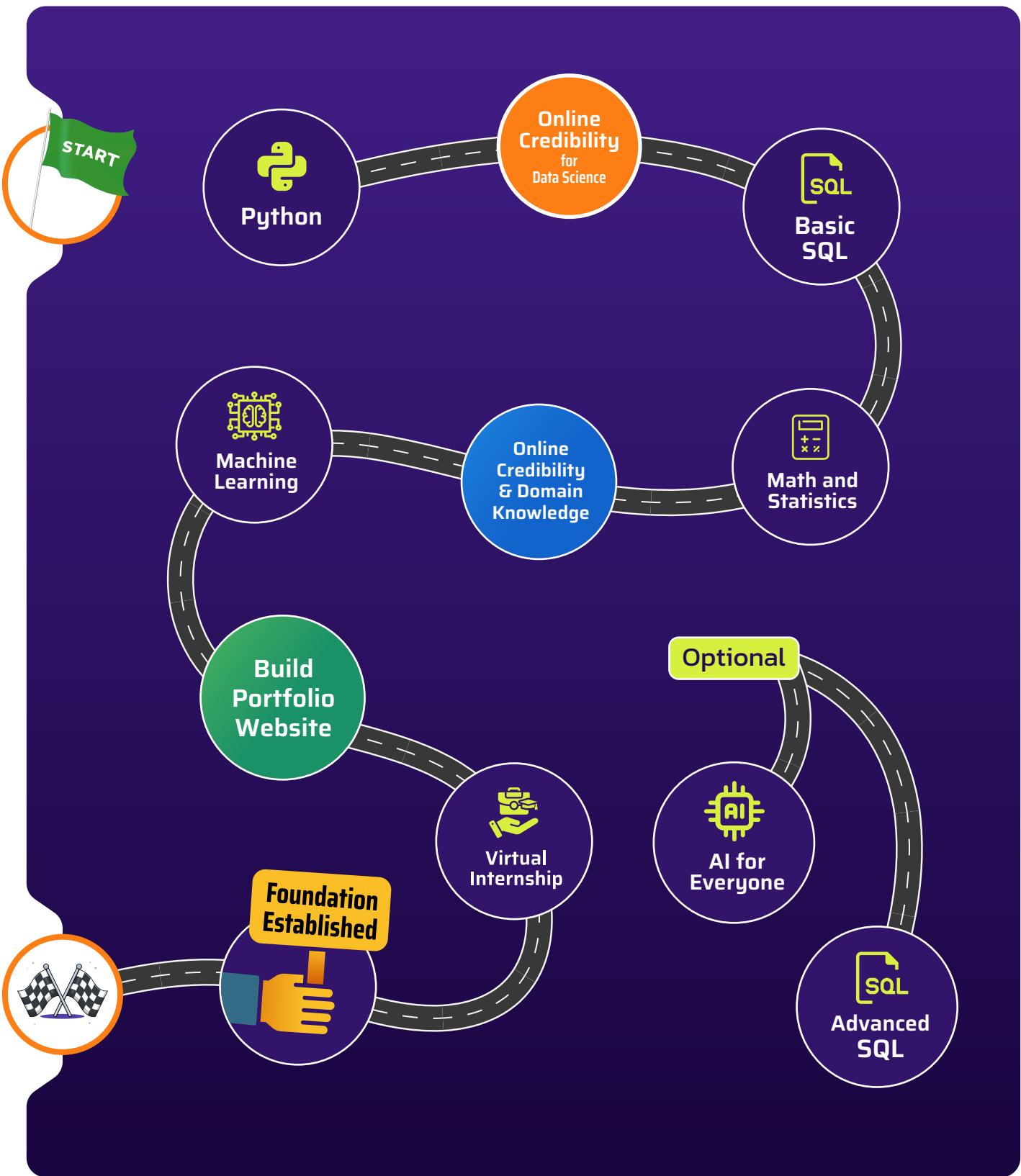
Get Job Ready



The course instructors being hiring managers for many years in big tech companies, they know what it takes to be truly job ready.

From resume-project challenges to building online credibility, we cover various aspects of job readiness. With practical exercises, industry-relevant content, our learners get access to different tools they need to get job-ready!

DATA SCIENCE FOUNDATION BOOTCAMP



WHY THIS IS THE MOST EFFECTIVE BOOTCAMP ON THIS PLANET



We offer unlimited chat support for doubt clearing through Discord.



The projects you will be working on will be related to Food & Beverages, Finance, Healthcare, Banking, Supply Chain Analytics and more will be added in the future.



We provide Virtual Internship, Project Portfolio Website and Building Online Credibility.



You will experience highly engaging content that provides a cinematic feel, real-world business practice problems, and interactive business meetings.



In this bootcamp, you won't be working on toy datasets. Instead, you'll tackle complex datasets used by real companies, with millions of records.



100% refund within 30 days and pay the difference feature.



PYTHON

BEGINNER TO ADVANCED FOR DATA PROFESSIONALS

This project-based learning course will teach Python through hands-on experience with two real-world projects: (1) Data analysis in the hospitality domain and (2) Building an expense tracking system. It is designed for both complete beginners with no prior coding experience and those with some Python knowledge looking to enhance their skills. The course is taught by an industry expert with over 14 years of experience at leading tech companies such as Nvidia and Bloomberg.



5710+
Learners

4.8
Rating

Screenshots

```

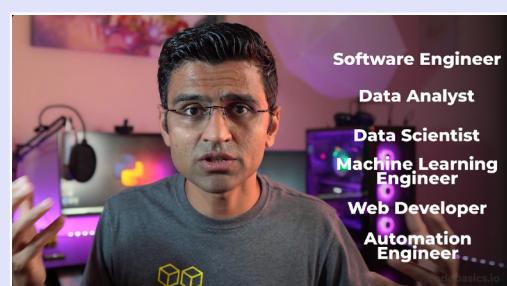
In [1]: movie_id      title          budget   revenue   year   studio_id   language_id
0    101  K.G.F.: Chapter 2  12.5 Millions  2422    8.4  Hornbill  5
1    102  Doctor Strange in the Multiverse of Madness  5622    7.0  Marvel Studios  5
2    103  Thor: The Dark World  5013    7.0  Marvel Studios  5
3    104  Iron Man 3  2817    7.0  Marvel Studios  5

In [2]: df_financials = pd.read_excel('movie_expenses.xlsx', 'Financials')
df_financials

```



codebasics.io



Software Engineer
Data Analyst
Data Scientist
Machine Learning Engineer
Web Developer
Automation Engineer



codebasics.io



RB Rhytham Bajai

Rupali Sherekar
Computer Science Professor
venturing into analytics

The entire course is wonderful. Everything is mapped to real world problems. Was specially excited to do the projects and I must say completing them was satisfying. Thanks Dhaval Patel Sir, you have very simple ways of explaining complex topics. Was fulfilling learning experience. Highly recommended for beginners as well as intermediate learners.

Python

Jun 04, 2024

★★★★★



urse
in
at



S Y Bharath
Data Analyst



It was a very helpful course that gave me a lot of information and hands-on experience, and I learned a lot about the basics and projects. It will be a good start to a career for anyone. I look forward to enrolling in a few more courses and upskilling myself more. Thank you.

Python

Jun 04, 2024

★★★★★

SC Shubham Chaurasia
Software Engineer



There are numerous Python tutorials available for beginners, but they often contain extensive content and lack live project implementations. However, this course was refreshingly concise. As a MEAN stack developer, I swiftly acquired the knowledge thanks to Mr. Dhaval, who made learning through this tutorial remarkably straightforward. The projects were based on real-world examples, making it ideal for those who prefer to learn by doing. I highly recommend enrolling in this course if you enjoy practical learning experiences.

Python

Jun 16, 2024

★★★★★

SH Shahlogue

Course delivers has promised. It is beyond the introduction to toolset, it's about developing the skills of data professional using the tools but also understanding the taught process of being an success data professional. Well structured and helpful, excellent job done here team codebasics.

Python

Apr 16, 2024

★★★★★



SYLLABUS

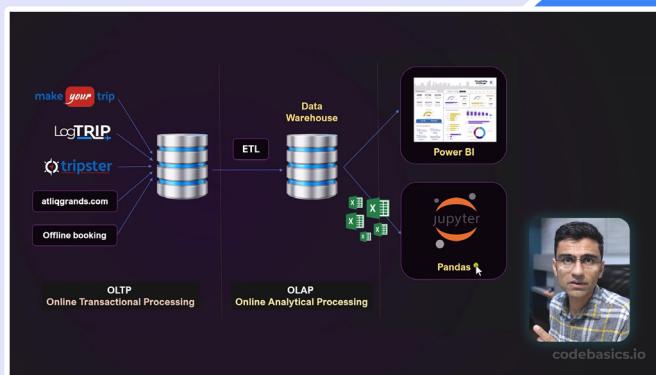
- | | |
|--|---|
| <ul style="list-style-type: none"><input checked="" type="checkbox"/> Python Installation - Windows<input checked="" type="checkbox"/> Python Installation - Linux<input checked="" type="checkbox"/> Python Installation - Mac<input checked="" type="checkbox"/> Variables<input checked="" type="checkbox"/> Numbers<input checked="" type="checkbox"/> Strings<input checked="" type="checkbox"/> Lists<input checked="" type="checkbox"/> Install Pycharm<input checked="" type="checkbox"/> If Condition<input checked="" type="checkbox"/> For Loop<input checked="" type="checkbox"/> Functions<input checked="" type="checkbox"/> Dictionary and Tuples<input checked="" type="checkbox"/> Modules and Pip<input checked="" type="checkbox"/> File Handling<input checked="" type="checkbox"/> Classes and Objects<input checked="" type="checkbox"/> Operator Overloading<input checked="" type="checkbox"/> Inheritance<input checked="" type="checkbox"/> Exception Handling<input checked="" type="checkbox"/> Main Function<input checked="" type="checkbox"/> Numpy Introduction and Benefits<input checked="" type="checkbox"/> Basic Operations<input checked="" type="checkbox"/> Matrix Operations<input checked="" type="checkbox"/> Slicing, Stacking<input checked="" type="checkbox"/> Pandas Introduction and Installation<input checked="" type="checkbox"/> Dataframe Basics | <ul style="list-style-type: none"><input checked="" type="checkbox"/> Read, Write Excel and CSV Files<input checked="" type="checkbox"/> Handle NA values<input checked="" type="checkbox"/> Group By<input checked="" type="checkbox"/> Concat and Merge<input checked="" type="checkbox"/> Data Visualization Using Matplotlib and Seaborn<input checked="" type="checkbox"/> Data Exploration<input checked="" type="checkbox"/> Data Cleaning<input checked="" type="checkbox"/> Data Transformation<input checked="" type="checkbox"/> Insights Generation<input checked="" type="checkbox"/> Set and Frozenset<input checked="" type="checkbox"/> Lists, Dict and Set Comprehensions<input checked="" type="checkbox"/> Code Debugging Using PyCharm<input checked="" type="checkbox"/> Working with JSON<input checked="" type="checkbox"/> Generators and Iterators<input checked="" type="checkbox"/> Decorators<input checked="" type="checkbox"/> API<input checked="" type="checkbox"/> Calling APIs with request Package<input checked="" type="checkbox"/> Building APIs with FastAPI<input checked="" type="checkbox"/> Automated Testing with Pytest<input checked="" type="checkbox"/> Data Validation with Pydantic<input checked="" type="checkbox"/> Database CRUD Operations<input checked="" type="checkbox"/> Streamlit Introduction (Frontend)<input checked="" type="checkbox"/> Analytics: Backend (FastAPI) & Frontend (Streamlit) |
|--|---|

BUSINESS PROJECTS

01

HOSPITALITY DOMAIN DATA ANALYSIS

PYTHON



```

Out[96]:
bookings_id property_id booking_date check_in_date checkout_date no_guests room_category booking_platform ratings_g
Int64Index: 134573 entries, 0 to 134572
Data columns (total 15 columns):
 #   Column          Non-Null Count  Dtype  
--- 
 0   booking_id      134573 non-null  object 
 1   property_id     134573 non-null  int64  
 2   booking_date    134573 non-null  object 
 3   check_in_date   134573 non-null  object 
 4   checkout_date   134573 non-null  object 
 5   no_guests       134573 non-null  float64
 6   room_category   134573 non-null  object 
 7   booking_platform 134573 non-null  object 
 8   ratings_given   56676 non-null  float64
 9   booking_status  134573 non-null  object 
 10  revenue_generated 134573 non-null  int64  
 11  revenue_realized 134573 non-null  int64  
 12  property_name   134573 non-null  object 
 13  category        134573 non-null  object 
 14  city            134573 non-null  object 
dtypes: float64(2), int64(3), object(10)
memory usage: 16.4+ MB

```

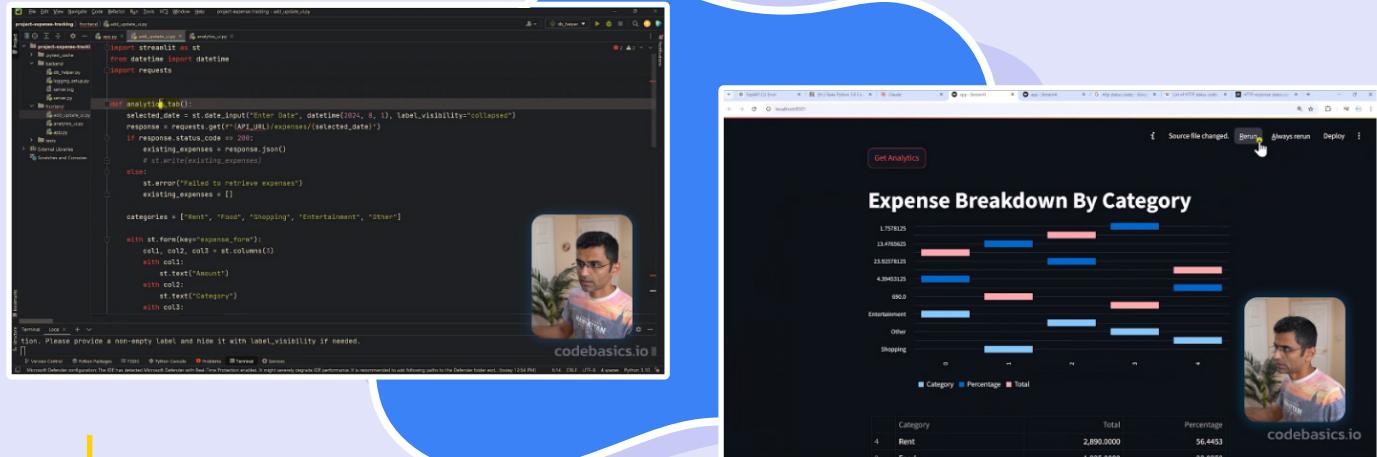
By completing this project, you will:

- Gain deep domain knowledge in the hotel industry, specifically in data analysis for operations, customer behaviour, revenue management, and market trends.
- Acquire essential technical skills using Python's Pandas library, including data cleaning, transformation, EDA, statistical analysis, and visualization.
- Recognize the high relevance of data-driven insights in today's business world for making informed decisions, optimizing operations, and enhancing customer experiences, not just in the hotel industry but across various sectors.

02

EXPENSE TRACKING SYSTEM

— PYTHON



By completing this project, you will:

- Develop practical knowledge of expense management systems by building an application that tracks and analyses day-to-day expenses, such as rent, shopping, and food, providing users with a comprehensive view of their spending habits.
- Master key technical skills, including Python basics, database interactions with MySQL, backend development using FastAPI, and frontend design using the Streamlit framework, gaining end-to-end application development experience.
- Understand the value of analytics in financial decision-making by implementing features like category-wise and month-wise expense breakdowns, making this project relevant for aspiring software engineers, data scientists, and data analysts.

SQL

BEGINNER TO ADVANCED FOR DATA PROFESSIONALS

Beginners to advanced course for those preparing for a data career (data analyst, data scientist, or data engineer). This course is carefully curated to **simulate real-time organizational experience** to prepare you for the current job market and at the same time provides you with an ultimate learning experience through a **storytelling mode** that you would see in movies.



7420+
Learners

4.9
Rating



AV Aman Verma

Dear Peter, I wanted to take a moment to express my sincerest gratitude for your invaluable assistance in helping me attain my SQL certificates. Your guidance and support throughout the process have been truly instrumental in my success. Specially i love the movies dataset your have created. I am also a big Marvel Fan!!!!!! :-)

Apr 08, 2024

★★★★★

AP Ashish Patel

The Codebasics SQL course is easy to understand and provides valuable insights into SQL. The mentor Dheval Patel breaks down complex concepts into simple, digestible lessons, making it accessible for beginners. Each module builds on the previous one, ensuring a strong foundational knowledge. Whether you're new to SQL or looking to solidify your skills, this course offers practical examples and clear explanations that make learning SQL straightforward and enjoyable.

May 28, 2024

★★★★★

SB Shrabani Guha Biswas

This SQL Course is totally different from the other platform courses and I totally loved the content that Peter Pandey has created. The real life scenarios and the challenges faced and the tips and tricks shared to handle those scenarios was awesome. Thanks for the guided tutorial and it will help me to crack interviews in the coming days as an experienced professional.

Apr 15, 2024

★★★★★

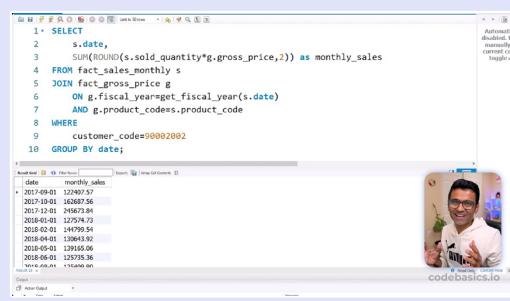
MB Menak Baskar

Hi Peter Pandey, I set my expectations very minimal before starting this course, but the way you taught was beyond anyone's expectations. Never felt bored in any of the modules, it was very interactive and easy to follow. Covered almost all topics. New ones or the experienced person - Both will like this course. I highly recommend these courses to my colleagues. Thanks a lot for your contribution and keep up doing good works :)

May 03, 2024

★★★★★

Screenshots



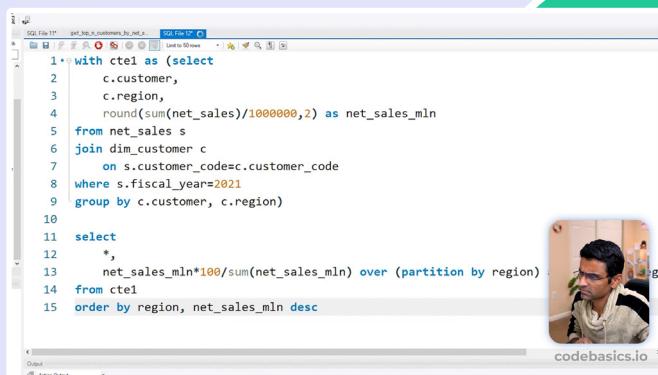
SYLLABUS

- Install MySQL: Windows
- Install MySQL: Linux, Mac
- Retrieve Data Using Text Query (SELECT, WHERE, DISTINCT, LIKE)
- Retrieve Data Using Numeric Query (BETWEEN, IN, ORDER BY, LIMIT, OFFSET)
- Summary Analytics (MIN, MAX, AVG, GROUP BY)
- HAVING Clause
- Calculated Columns (IF, CASE, YEAR, CURYEAR)
- Joins (INNER, LEFT, RIGHT, FULL,CROSS)
- Join More Than Two Tables
- Subqueries
- ANY, ALL Operators
- Correlated Subquery
- Common Table Expression (CTE)
- CTE Benefits & Other Applications
- Entity Relationship Diagram (ERD)
- Database Normalization and Data Integrity
- Data Types: Numeric (INT, DECIMAL, FLOAT, DOUBLE)
- Data Types: String (VARCHAR, CHAR, ENUM)
- Data Types: Date, Time (DATETIME, DATE, TIME, YEAR, TIMESTAMP)
- Data Types: JSON, Spatial (JSON, GEOMETRY)
- Primary Key
- Foreign Key
- Entity Relationship Diagram – ERD
- INSERT, UPDATE, DELETE Statement
- ETL, Data Warehouse
- OLAP vs OLTP, Data Catalog
- Fact vs Dimension Table
- Star vs Snowflake Schema
- Data Import
- Simplified: What is Kanban?
- User-Defined Functions
- Stored Procedures
- Database Views
- Window Functions OVER Clause, ROW_NUMBER, RANK, DENSE_RANK
- Database Triggers
- Database Events
- User Accounts and Privileges
- Database Indexes

BUSINESS PROJECTS

03 FINANCE AND TOP N INSIGHTS

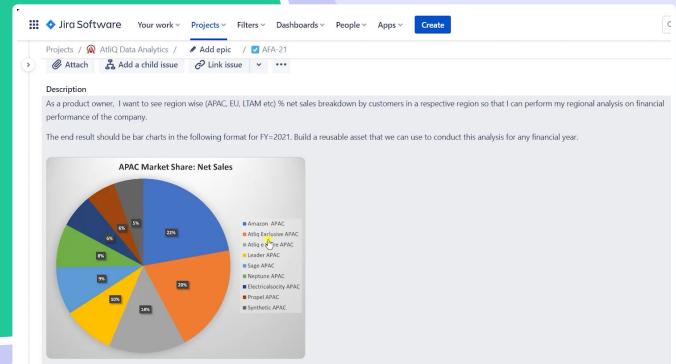
— SQL



```

1+ with cte1 as (select
2+     c.customer,
3+     c.region,
4+     round((sum(net_sales)/1000000,2) as net_sales_mln
5+ from net_sales s
6+ join dim_customer c
7+     on s.customer_code=c.customer_code
8+ where s.fiscal_year=2021
9+ group by c.customer, c.region)
10+
11 select
12     *,
13     net_sales_mln*100/sum(net_sales_mln) over (partition by region)
14     from cte1
15     order by region, net_sales_mln desc

```



- Gain domain knowledge in finance analytics for the consumer goods industry, focusing on generating reports for key metrics.
- Develop technical skills in efficient SQL querying, utilizing User-Defined Functions, creating and executing stored procedures, and leveraging views for data analysis.
- Understand the significance of data-driven financial decision-making in today's competitive business landscape.
- Drive sustainable growth by utilizing automated reporting solutions, such as stored procedures, to provide accurate and real-time information for financial analysis and decision-making.

04 | SUPPLY CHAIN ANALYTICS AND MODEL OPTIMISATION

SQL



```

1 • with cte
2 o   SELECT
3     s.customer_code,
4     sum(s.sold_quantity) as total_sold_qty,
5     sum(s.forecast_quantity) as total_forecast_qty,
6     sum((forecast_quantity-sold_quantity)) as net_err,
7     sum((forecast_quantity-sold_quantity))*100/sum(forecast_quantity) as net_err_pct,
8     sum(abs(forecast_quantity-sold_quantity)) as abs_err,
9     sum(abs(forecast_quantity-sold_quantity))*100/sum(forecast_quantity) as abs_err_pct
10    from gdb041.fact_act_est s
11   where s.fiscal_year=2021
12   group by customer_code;
13 • select
14   e.*,
15   c.customer,
16   c.market,
17   if (abs_err_pct > 100, 0, 100-abs_err_pct) as forecast_accuracy
18  from forecast_err_table e
19  join dim_customer c

```

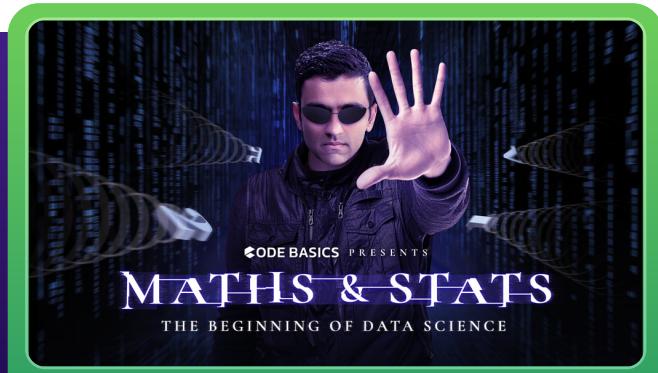
By completing this project, you will:

- Gain domain knowledge in supply chain analytics for the consumer goods industry, focusing on forecasting errors and absolute error calculation.
- Develop technical skills in creating helper tables, managing triggers, and optimizing query execution time.
- Understand the significance of accurate demand forecasting for consumer goods companies and how supply chain analytics can achieve it.
- Acquire practical experience through hands-on exercises and real-world examples in consumer goods supply chain analytics.

MATH AND STATISTICS

FOR AI, DATA SCIENCE

Learn the key concepts of Math and Statistics that lay the foundations for a strong data science career. This course is carefully curated to simulate real-time organizational experience to prepare you for the current job market and at the same time provides you with an ultimate learning experience through storytelling and intuitive explanations.



1150+
Learners

5.0
Rating



Noble Sunil



Codebasics "Math and Statistics for Data Science" course is a hidden gem! Despite not taking math in 11th and 12th grade, I was able to grasp every concept effortlessly. This course gave me a fresh perspective on data science, and the best part? It's incredibly affordable. Whether you're a beginner or brushing up on your skills, this course is a must!

Math & Stats

Sep 18, 2024

★★★★★

NS Naraharisetti Dhanesh
Naga Surya



This course is an excellent starting point for aspiring data scientists. Its clear and concise explanations make complex concepts approachable, inspiring you to stay motivated and excited about your journey towards becoming a data scientist. I'm eager to dive deeper into the machine learning course and data science bootcamp.

Math & Stats

Aug 21, 2024

★★★★★

PO Poojasree



I have thoroughly enjoyed the course; it is intuitive and well-structured. The course not only covers theoretical concepts but also provides real-world examples of how each statistical topic is applied in the industry. This approach ensures a solid foundation in math and statistics. Codebasics did an excellent job in crafting the course, and Dhaval sir taught it exceptionally well.

Math & Stats

Jul 14, 2024

★★★★★

Piyush
Student



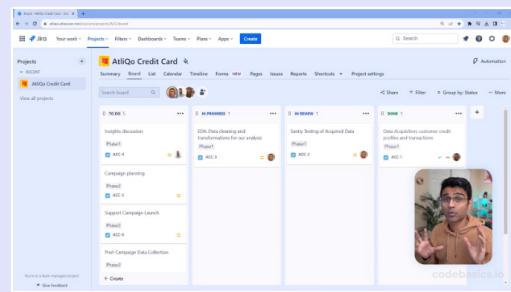
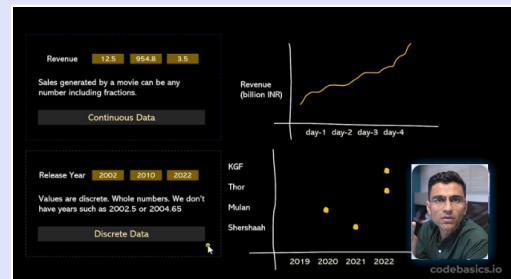
The way Dhaval Patel and whole codebasics team explained each topic with practical implementations is very Amazing about this course. I liked that this course covered only those topics which are going to be helpful in future and which are actually needed, comparing to those course which include all topics and which creates a fear of learning. Grateful to the whole codebasics team.

Math & Stats

May 02, 2024

★★★★★

Screenshots



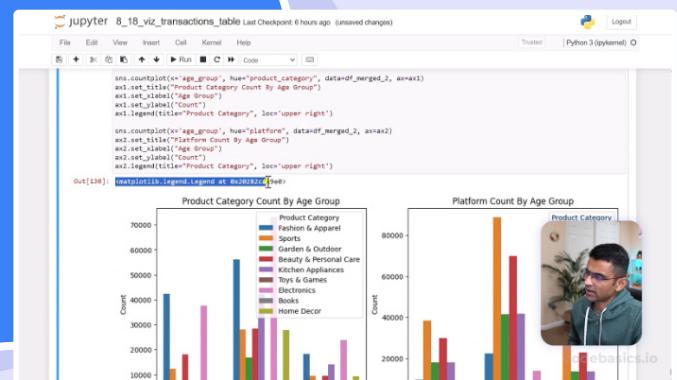
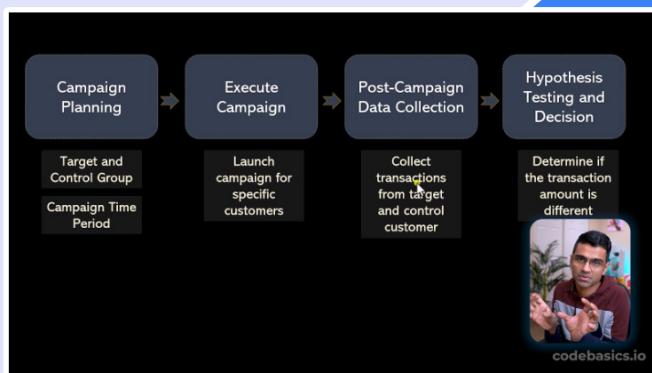
SYLLABUS

- | | |
|--|--|
| <ul style="list-style-type: none"><input checked="" type="checkbox"/> Types of Data<input checked="" type="checkbox"/> Pie Chart and Bar Chart<input checked="" type="checkbox"/> Histograms and Line Chart<input checked="" type="checkbox"/> Scatter and Bubble Plot<input checked="" type="checkbox"/> Univariate vs. Bivariate vs. Multivariate Analysis<input checked="" type="checkbox"/> Numpy Introduction and Benefits<input checked="" type="checkbox"/> Basic Operations<input checked="" type="checkbox"/> Matrix Operations<input checked="" type="checkbox"/> Slicing, Stacking<input checked="" type="checkbox"/> Descriptive vs. Inferential Statistics<input checked="" type="checkbox"/> Measures of Central Tendency: Mean, Median, Mode<input checked="" type="checkbox"/> Percentile<input checked="" type="checkbox"/> Measures of Dispersion: Range, IQR<input checked="" type="checkbox"/> Box or Whisker Plot<input checked="" type="checkbox"/> Outlier Treatment Using IQR and Box Plot<input checked="" type="checkbox"/> Measures of Dispersion: Variance and Standard Deviation<input checked="" type="checkbox"/> Correlation vs Causation<input checked="" type="checkbox"/> Probability Basics<input checked="" type="checkbox"/> Addition and Multiplication Rule<input checked="" type="checkbox"/> Conditional Probability and Bayes Theorem<input checked="" type="checkbox"/> Distribution | <ul style="list-style-type: none"><input checked="" type="checkbox"/> Skewness<input checked="" type="checkbox"/> Normal Distribution<input checked="" type="checkbox"/> Detect Outliers Using Normal Distribution<input checked="" type="checkbox"/> Z Score<input checked="" type="checkbox"/> Standard Normal Distribution (SND)<input checked="" type="checkbox"/> Kanban Project Management<input checked="" type="checkbox"/> Random Sampling & Sample Bias<input checked="" type="checkbox"/> The Law of Large Numbers<input checked="" type="checkbox"/> Central Limit Theorem, Sampling Distribution<input checked="" type="checkbox"/> Standard Error<input checked="" type="checkbox"/> Confidence Interval<input checked="" type="checkbox"/> Null vs Alternate Hypothesis<input checked="" type="checkbox"/> Z Test, Rejection Region<input checked="" type="checkbox"/> p-Value<input checked="" type="checkbox"/> Statistical Power & Effect Size<input checked="" type="checkbox"/> A/B Testing<input checked="" type="checkbox"/> A/B Testing Using Z Test<input checked="" type="checkbox"/> T-test<input checked="" type="checkbox"/> Chi-squared Distribution<input checked="" type="checkbox"/> Chi-squared Test of Goodness of Fit<input checked="" type="checkbox"/> Chi-squared Test of Independence |
|--|--|

BUSINESS PROJECTS

05 ATLIQO BANK PROJECT

MATH AND STATISTICS



By completing this project, you will:

- Gain hands-on experience in the banking sector by working on a large-scale data analysis project aimed at identifying the target market for a new credit card launch, simulating real-life industry challenges with over 50,000 records.
- Master essential data analysis techniques, including data validation, cleaning, handling outliers, MySQL setup, and data visualization using Python, while performing critical tasks like credit score analysis, correlation studies, and hypothesis testing.
- Understand the role of A/B testing and statistical decision-making in marketing campaigns, providing insights into customer segmentation and campaign effectiveness, which are applicable across various industries.

MASTER MACHINE LEARNING FOR DATA SCIENCE & AI

BEGINNER TO ADVANCED

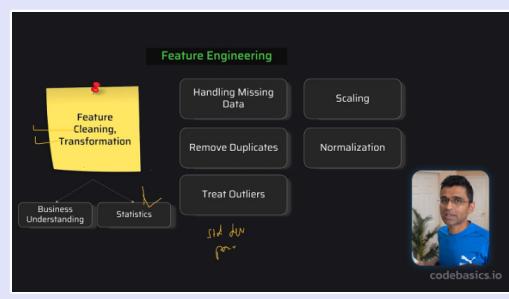
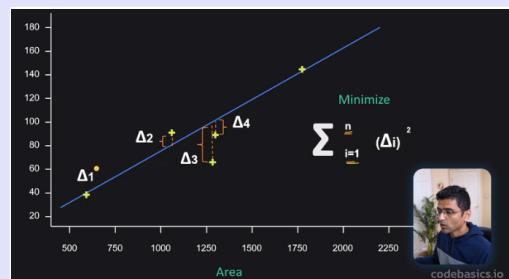
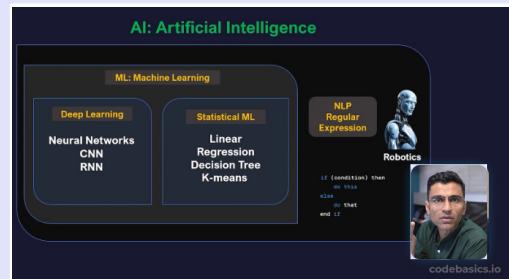
Learn Machine Learning from an AI entrepreneur with extensive industry experience and a popular YouTube channel (Codebasics) with 1 million subscribers. This course takes you from beginner to advanced levels, providing deep intuition on algorithms, engaging cinematic experiences, end-to-end projects, and hands-on coding practice. Designed for easy understanding, even for high school students, all at an affordable price.



1430+
Learners

5.0
Rating

Screenshots



Sandeep Jallu
Analyst



This is one of the best course in machine learning. The course is designed with beginners in mind, making it accessible even if you have little to no prior experience with machine learning. The explanation and clear and concise, with a strong focus on practical implementation rather just theory. It's a fantastic resource for building a strong foundation in machine learning.

Machine Learning

Sep 04, 2024



NS Naraharisetti Dhanesh Naga Surya



I have loved it .Their projects are incredibly relevant to the industry, providing hands-on experience that truly prepares students for the workforce. The Discord server is a fantastic resource for getting questions answered quickly. The mentors and other students are always there to offer help and support, making it a great place to learn and grow.

Machine Learning

Aug 21, 2024



GP Gopi Krishna Pujari

Thank you Dhaval Patel sir for this amazing course. Your clear explanations, engaging teaching style, and valuable insights made the learning experience truly enjoyable. Thank you for your dedication and commitment to helping students succeed..

Machine Learning

Aug 31, 2024



AM Aneesh Mohanan

I highly recommend this course to anyone interested in machine learning. It has equipped me with the knowledge and skills needed to tackle challenges and apply machine learning effectively. Special thanks to the instructors for their clear explanations and engaging teaching style.

Machine Learning

Sep 02, 2024



SYLLABUS

- Introduction to Machine Learning
- Classification vs Regression
- Supervised vs Unsupervised Learning
- Simple Linear Regression
- Multiple Linear Regression
- Cost Function
- Derivatives and Partial Derivatives
- Chain Rule
- Gradient Descent
- Model Evaluation
- Data Preprocessing: One Hot Encoding
- Polynomial Regression
- Overfitting and Underfitting
- L1 and L2 Regularization
- Bias Variance Trade Off
- Introduction to Classification
- Logistic Regression:
Binary Classification
- Model Evaluation: Accuracy, Precision
and Recall, F1 Score, Confusion Matrix
- Logistic Regression:
Multiclass Classification
- Cost Function: Log Loss
- Support Vector Machine (SVM)
- Data Pre-processing: Scaling
- Sklearn Pipeline
- Naive Bayes
- Decision Tree
- Handle Class Imbalance
- Introduction to Ensemble Learning
- Majority Voting, Average and
Weighted Average
- Bagging
- Random Forest
- Boosting: AdaBoost
- Gradient Boosting
- XGBoost
- Model Evaluation: ROC Curve & AUC
- Cost Benefit Analysis Using ROC
in Sklearn
- K Fold Cross Validation
- Stratified K Fold Cross Validation
- Hyperparameter Tuning
- 10 Stages of AI Project Lifecycle
- Requirements and Scope of Work (SOW)
- Data Collection, Data Cleaning &
Exploratory Data Analysis
- Feature Engineering
- Model Selection & Training
- Model Fine Tuning
- Model Deployment
- Monitoring and Feedback Using ML Ops
- Variance Inflation Factor (VIF)
- K Means Clustering

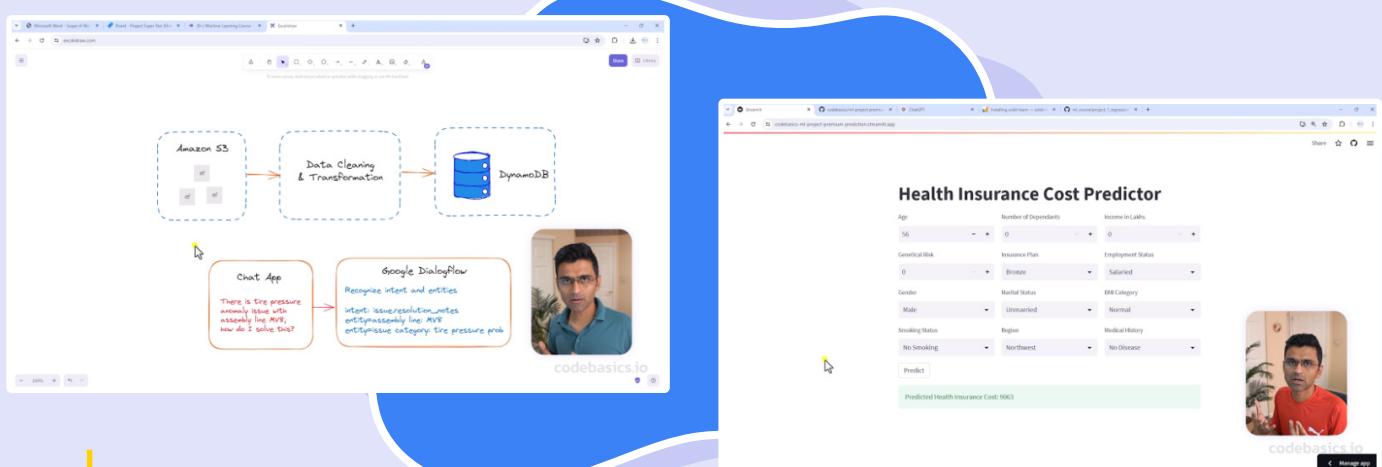
SYLLABUS

- Hierarchical Clustering
- DBSCAN
- ML Ops
- ML Flow
- Introduction & Basics of API
- Introduction to ML Cloud Platforms
- AWS Sagemaker
- Data Drift Detection Using PSI & CSI
- PSI & CSI: Practical Implementation

BUSINESS PROJECTS

06 HEALTHCARE PREMIUM PREDICTION (REGRESSION)

MACHINE
LEARNING

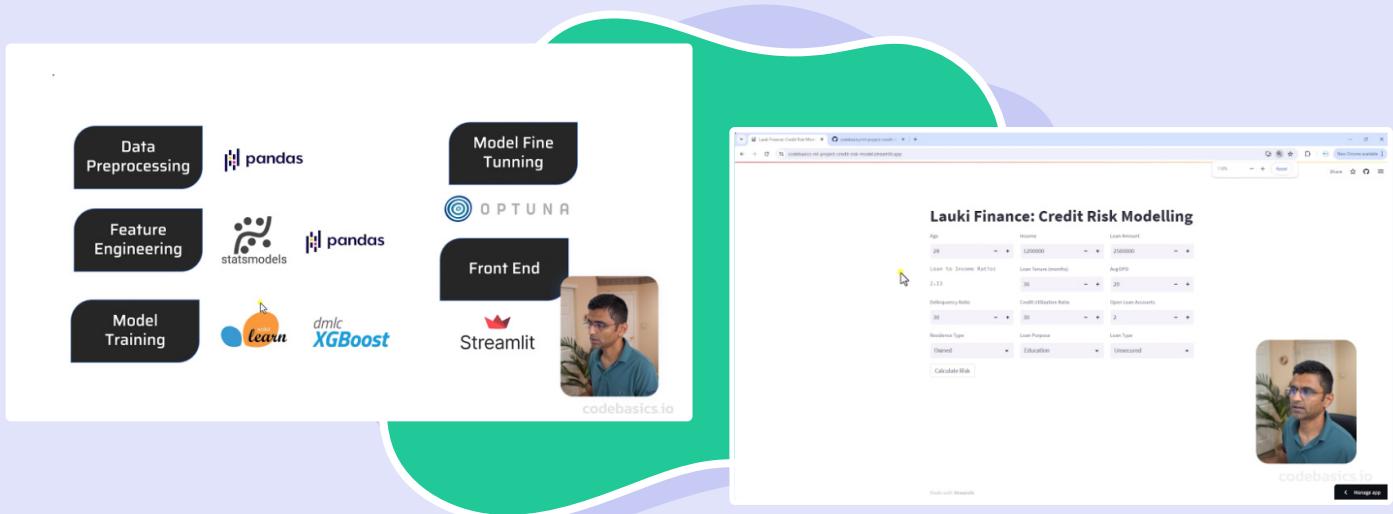


By completing this project, you will:

- Gain domain knowledge in the healthcare industry by predicting insurance premiums using machine learning regression models, covering real-world tasks like data collection, model training, and deployment.
- Develop advanced technical skills in data cleaning, exploratory data analysis (EDA), feature engineering, and regression model fine-tuning, achieving high accuracy while performing error analysis and segmentation.
- Understand the practical application of machine learning in predicting financial outcomes, and learning to build and deploy models using Python, Streamlit, and other tools for real-time decision-making across industries.

07 CREDIT RISK MODELLING (CLASSIFICATION)

MACHINE LEARNING



By completing this project, you will:

- Gain valuable insights into the financial domain, particularly non-banking financial companies (NBFCs) and credit approval processes, by developing a credit risk model to classify potential defaulters.
- Acquire technical expertise in data cleaning, exploratory data analysis (EDA), feature engineering (WOE/IV), and model training, with a focus on classification techniques, model fine-tuning using Optuna, and evaluation metrics like KS Statistic and Gini Coefficient.
- Learn how to deploy machine learning models by building a Streamlit app, delivering business presentations, and applying credit risk modelling principles across various financial sectors.

VIRTUAL INTERNSHIP

In our virtual Internship, you'll gain practical experience in Data Science foundations. You'll gain good knowledge in data cleaning, data validation, reporting, feature engineering, model deployment and client handling. You will also learn scrum project management and improve debugging skills. With a problem-solving mindset, you'll be ready to tackle any data challenge and make a meaningful impact.



Data Science Internship Offer

Dear Kirandeep,

We are pleased to inform you that you have been selected as a **Data Science Intern** at our organization. We welcome you to our team and are excited to have you onboard!

As a Data Science Intern, you will collaborate with our team of experienced professionals to develop predictive models and analyze data to help us make informed business decisions. Your work will be essential to our organization's success, and we are confident that you will make valuable contributions during your internship.

The internship will be conducted virtually, and we will provide you with all the necessary tools and resources to help you excel in your role. Please do not hesitate to reach out to us if you have any questions or concerns. We are here to support you and ensure that your virtual internship is a success.

Once again, congratulations on being selected for the Data Science Intern position. We look forward to working with you and wish you all the best during your internship.

Best regards,



Abhilasha Kashyap
HR Specialist | AtliQ Technologies Pvt. Ltd.
Website: www.atliq.com

CodeX Project: Streamlit App Development for Client Demo

Kirandeep!

We're now moving forward with an important step: **developing a Streamlit app**. Your task is to create an interactive app using the best-performing model from the modeling phase.

This app will be used to demo our pricing prediction solution to the client. A well-designed app will greatly impress the client and give them confidence in the solution we've developed.

To help you visualize the final product, I've attached a **demo file** that provides an idea of how the app should look and function. Please use it as a reference while building the Streamlit app.

Let me know if you have any questions during the development process, and I look forward to seeing the final product!

Best regards,



Hem
Head of AI/Data Science | AtliQ Technologies Pvt. Ltd.
Website: www.atliq.com

Customer Behavior Analysis

Kirandeep,

Following up on our call, our client, an e-commerce company, has provided us with a dataset containing various metrics tracking customer behavior, purchasing patterns, and cross-sell conversion rates. Your task is to analyze this data using **mathematical and statistical** concepts to provide insights that will help improve their marketing strategies.

The client has requested answers to key business questions. You may use any tool, such as Pandas, NumPy, statistical libraries like SciPy, Excel, etc., to complete the task.

Click on the "Download Files" button above to access all the files. Your specific task is described in '**client_requests.pdf**'. Be sure to check '**meta_data.txt**' for a thorough understanding of the dataset.

Feel free to ask your seniors for any questions!

Good luck with this task!

Best regards,



Hem
Head of AI/Data Science | AtliQ Technologies Pvt. Ltd.
Website: www.atliq.com



BUSINESS PROJECTS

08 BEVERAGE PRICE RANGE PREDICTION

CodeX Beverage: Price Prediction

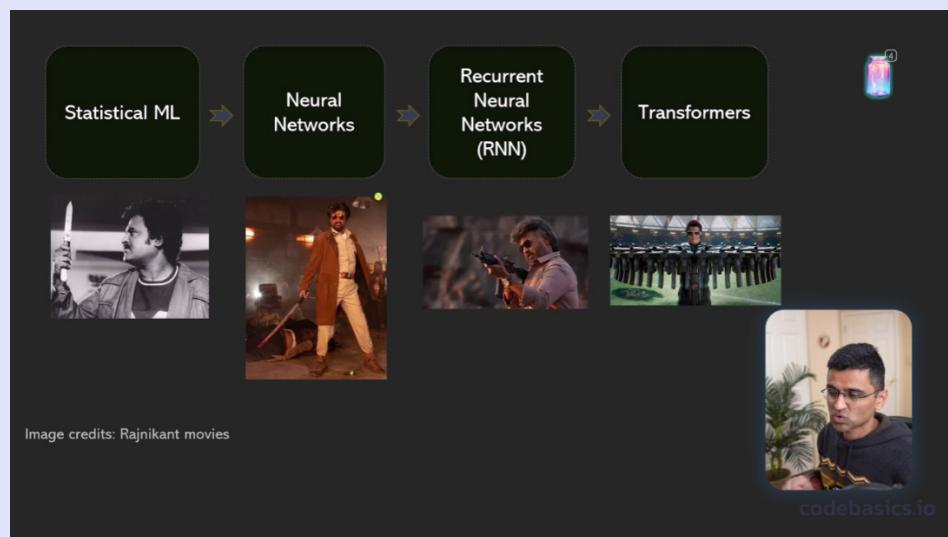
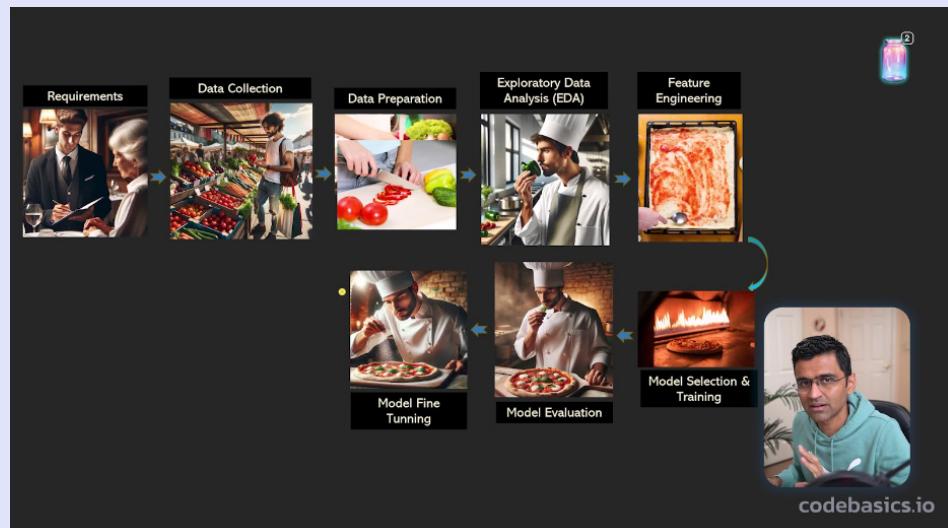
Age 30	Gender M	Zone Metro	Occupation Entrepreneur
Income Level (In L) 16L - 25L	Consume Frequency(weekly) 5-7 times	Current Brand Established	Preferable Consumption Size Medium (500 ml)
Awareness of other brands 2 to 4	Reasons for choosing brands Quality	Flavor Preference Traditional	Purchase Channel Online
Packaging Preference Premium	Health Concerns Medium (Moderately health-conscious)	Typical Consumption Situations Active (e.g. Sports, Gym)	
<input type="button" value="Calculate Price Range"/> Price Range: 200-250 INR			

By completing this project, you will:

- Gain valuable insights into market pricing strategies by building a predictive model that forecasts optimal price ranges for beverages, helping businesses avoid overselling or underselling based on survey data.
- Develop essential technical skills in machine learning, including data cleaning, validation, feature engineering, and model deployment using Python, ensuring the creation of an accurate and reliable classification model.
- Understand how predictive models drive business growth and competitiveness by enabling data-driven pricing decisions, applicable across various sectors beyond the beverage industry.

AI FOR EVERYONE

"AI For Everyone" is an optional module designed as a follow-up to our Data Science Foundation Bootcamp. In this module, we'll explore AI topics like Deep Learning, Gen AI, NLP, and how AI is used in different industries. We won't be doing any coding in this module. Instead, we'll focus on giving you a solid understanding of these concepts. Coding will come later in our final bootcamp, where we'll dive deep into these topics. This module is your chance to get comfortable with AI without the pressure of coding—perfect for starting your journey into the more technical side later!



FAQs

01 What prior knowledge or skills do I need before enrolling in this bootcamp?

The only thing you need as a prerequisite is the passion and interest in learning AI and Data Science. Actual prerequisites such as Python, Math, and Statistics basics are all covered in this course.

02 How much coding/programming do I need to know to become a Data Scientist?

The first module in our bootcamp is the Python module where we teach everything from basics. You don't need to be an expert coder, but proficiency in data manipulation, visualization, and basic algorithms is essential.

03 What if I have an unrelated degree but want to pursue data science?

A computer science degree or an IT degree is not required. We have many success stories of folks from non-tech backgrounds successfully cracking into a data science career.

04 Can I enroll in the bootcamp if I have no background in mathematics or statistics?

Yes. In this bootcamp, we have a module on Math and statistics. This module designed for absolute beginners, simplifies learning Math and Statistics through project-based learning with a real dataset, focusing on practical data science applications.

05 What kind of support is available if I have questions during the bootcamp?

We have created every lecture of this bootcamp with the motive to explain everything in an easy-to-understand manner. While working on these lectures you could make mistakes in steps or have some doubts. You need to commit yourself to hold patience, make efforts & figure out the solution by googling in order to become truly job ready.

For any questions, that Google cannot answer or if you hit a wall - we got you covered!

You can join our active discord community, which is a dedicated platform to discuss & clear your doubts with fellow learners & mentors.

FAQs

06 Is there a community platform where I can interact with other students?

Yes, you'll be part of an active Discord community where you can interact, and learn from one another.

07 Are the lectures going to be LIVE?

To keep the bootcamp affordable, we have kept the lectures as self-paced videos that you can watch at your convenience. However, there will be monthly seminars and live discord community support.

08 Do you provide any Virtual Internship?

Yes, there is Virtual Internship, included in this BootCamp.

09 Can I expect to work on real-world datasets during the course?

Yes. In this bootcamp, we have a module on Math and statistics. This module designed for absolute beginners, simplifies learning Math and Statistics through project-based learning with a real dataset, focusing on practical data science applications.

10 Does this bootcamp provide lifetime access to the content?

Yes, this bootcamp will be available to you for a lifetime.

11 Will I receive a certificate upon completing the bootcamp?

Upon successfully completing each tool, you will receive a certificate. In total, you'll earn four individual certificates for Python, SQL, Maths & Statistics, and Machine Learning, as well as another one for completing the entire Bootcamp. Additionally, you'll also receive a letter of completion for the Virtual Internship.

12 How do I get the certificate?

You need to complete at least 70% of the course and score 70% or more on the final quiz. After completing the quiz, simply click "complete and get certificate" to download your certificate.

FAQs

13 Do you have an EMI (Installment) option?

Our website doesn't support an EMI option right now. But don't worry, you can still pay in installments by following these steps,

- 1 First buy the Python course.
- 2 After you have completed the Python course, buy the Math and Statistics course.
- 3 After completing the Math and Statistics course, buy the Bootcamp.

This way you will pay in 3 installments

Remember that the Bootcamp includes the same content from individual Python and Math and Statistics courses that you bought previously hence you will not lose anything in terms of learning in your Bootcamp curriculum.

Also, all the progress you made in these individual courses will be transferred to Bootcamp automatically."

14 Will I receive a certificate upon completing the bootcamp?

Upon successfully completing each tool, you will receive a certificate. In total, you'll earn four individual certificates for Python, SQL, Maths & Statistics, and Machine Learning, as well as another one for completing the entire Bootcamp. Additionally, you'll also receive a letter of completion for the Virtual Internship.

15 I have already purchased other Codebasics courses like SQL and Python. Will I have to pay the full amount to enroll in this bootcamp?

No, since both of these courses are part of this bootcamp, the amount you've already paid for the SQL and Python courses will be deducted from the bootcamp price.

16 I'm already enrolled in the Codebasics Data Analytics Bootcamp. Will I have to pay the full amount to enroll in this bootcamp?

No, since the SQL and Python courses in the Data Analytics Bootcamp are also part of this bootcamp, the price you've paid for those courses will be deducted from the bootcamp price.

FAQs

17 What if I don't like this bootcamp? Is there a refund policy?

Yes, we offer a no-questions-asked refund policy. No hidden conditions—if you request a refund within 30 days of enrollment, you will receive a 100% refund.

18 How can I request a refund?

Simply email us at info@codebasics.io with your name, course name, course purchase date, and the registered email ID you used on the Codebasics website. We will initiate the refund after verifying your details and eligibility according to our refund policy.

FROM FOUNDER'S DESK

Hello, I'm Dhaval Patel, and I want to share how and why I started the Codebasics YouTube channel. It all began during one of the darkest periods of my life. In **2011**, while working for Bloomberg, I was diagnosed with **Ulcerative colitis**, a painful autoimmune disease. Despite trying numerous treatments, my health continued to decline. By **2015**, I felt hopeless and **constantly in pain**.



During this time, I stumbled upon a life-changing quote:

"Life is 10% what happens to you and 90% how you react."

Determined to shift my focus and find joy, I sought an activity that I enjoyed and that allowed me to help others. **Given my teaching abilities and inclination to share knowledge, I started the Codebasics YouTube channel.** Teaching became an outlet for me, enabling me to make a positive impact and share my expertise.

With each "Thank you" comment I received on my videos, my sense of purpose and will to live grew stronger. Over time, the Codebasics community flourished, and we now have over **1.11 Million+ subscribers** on YouTube and more than **35K+ paid learners**. I've left my full-time job in the year 2022 to dedicate myself to Codebasics fully. My goal is to continue sharing my knowledge and creating a **positive impact** through this platform.

Use my story as to remind yourself that life is not always going to be easy but you can rise above that and make the world a better place to live for yourself and others.



121K+



34K+



1.12M+



441K+



info@codebasics.io



www.codebasics.io