UDACITY
FOR ENTERPRISE

Data Analyst

Data Analyst Nanodegree Program

BUILT IN COLLABORATION WITH

kaggle

This program prepares you for a career as a data analyst by helping you learn to organize data, uncover patterns and insights, draw meaningful conclusions, and clearly communicate critical findings. You'll develop proficiency in Python and its data analysis libraries (Numpy, pandas, Matplotlib) and SQL as you build a portfolio of projects to showcase in your job search.

Depending on how quickly you work through the material, the amount of time required is variable. We have included an hourly estimation for each section of the program. In order to succeed in this program, we recommend having experience working with data in Python (NumPy and Pandas) and SQL.

Program Information

TIME
4 months
Study 10 hours/week

LEVEL
Practitioner

PREREQUISITES
experience working with data in Python (specifically Numpy and Pandas) and SQL. Including Python standard libraries, and working with data with Pandas and Numpy.

HARDWARE/SOFTWARE REQUIRED
Access to the Internet, and a 64 bit computer. Additional software such as Python and its common data analysis libraries (e.g., Numpy and Pandas) will be required, but the program will guide learners on how to download once the course has begun.

LEARN MORE ABOUT THIS NANODEGREE
Contact us at enterpriseNDs@ udacity.com.

Our Classroom Experience
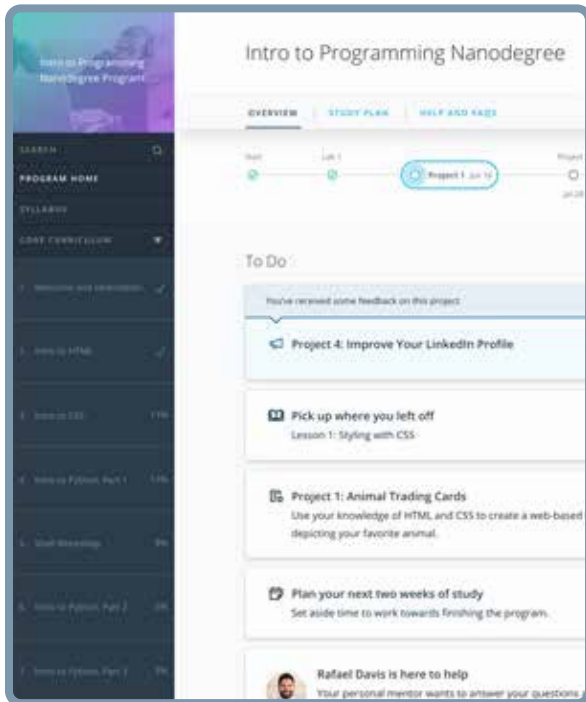
REAL-WORLD PROJECTS
Learners build new skills through industry-relevant projects and receive personalized feedback from our network of 900+ project reviewers. Our simple user interface makes it easy to submit projects as often as needed and receive unlimited feedback.

KNOWLEDGE
Answers to most questions can be found with Knowledge, our proprietary wiki. Learners can search questions asked by others and discover in real-time how to solve challenges.

LEARNER HUB
Learners leverage the power of community through a simple, yet powerful chat interface built within the classroom. Learner Hub connects learners with their technical mentor and fellow learners.

WORKSPACES
Learners can check the output and quality of their code by testing it on interactive workspaces that are integrated into the classroom.

QUIZZES
Understanding concepts learned during lessons is made simple with auto-graded quizzes. Learners can easily go back and brush up on concepts at anytime during the course.
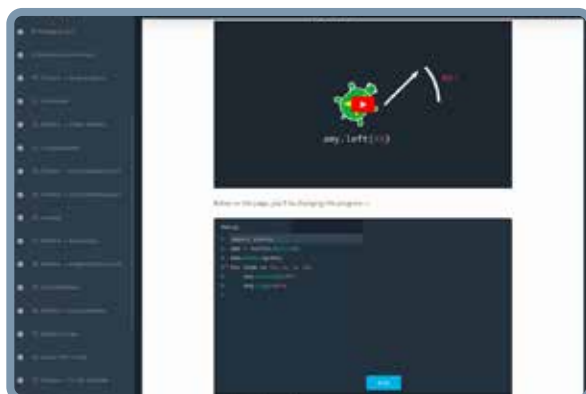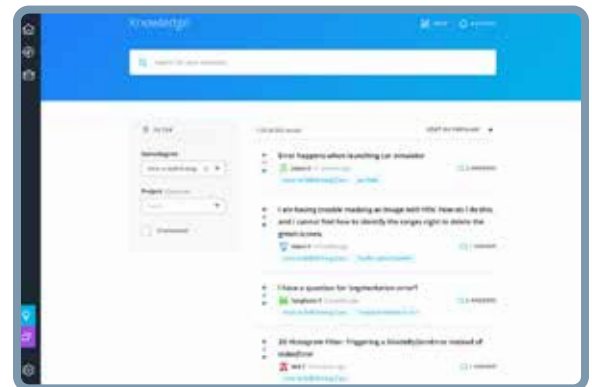
CUSTOM STUDY PLANS
Mentors create a custom study plan tailored to learners' needs. This plan keeps track of progress toward learner goals.

PROGRESS TRACKER
Personalized milestone reminders help learners stay on track and focused as they work to complete their Nanodegree program.

Learn with the Best

Josh Bernhard
DATA SCIENTIST AT NERD WALLET

Josh has been sharing his passion for data for nearly a decade at all levels of university, and as Lead Data Science Instructor at Galvanize. He's used data science for work ranging from cancer research to process automation.

Sebastian Thrun
INSTRUCTOR

As the founder and president of Udacity, Sebastian's mission is to democratize education. He is also the founder of Google X, where he led projects including the Self-Driving Car, Google Glass, and more.

Derek Steer
CEO AT MODE

Derek is the CEO of Mode Analytics. He developed an analytical foundation at Facebook and Yammer and is passionate about sharing it with future analysts. He authored SQL School and is a mentor at Insight Data Science.

Juno Lee
CURRICULUM LEAD AT UDACITY

Juno is the curriculum lead for the School of Data Science. She has been sharing her passion for data and teaching, building several courses at Udacity. As a data scientist, she built recommendation engines, computer vision and NLP models, and tools to analyze user behavior.

Mike Yi
DATA ANALYST INSTRUCTOR

Mike is a Content Developer with a multidisciplinary academic background, including math, statistics, physics, and psychology. Previously, he worked on Udacity's Data Analyst Nanodegree program as a support lead.

David Venturi
DATA ANALYST INSTRUCTOR

Formerly a chemical engineer and data analyst, David created a personalized data science master's program using online resources. He has studied hundreds of online courses and is excited to bring the best to Udacity students.

Sam Nelson
PRODUCT LEAD

Sam is the Product Lead for Udacity's Data Analyst, Business Analyst, and Data Foundations programs. He's worked as an analytics consultant on projects in several industries, and is passionate about helping others improve their data skills.

Course 1: Introduction to Data Analysis

Learn the data analysis process of questioning, wrangling, exploring, analyzing, and communicating data. Learn how to work with data in Python using libraries like NumPy and Pandas.

Project	Explore Weather Trends

This project will introduce you to the SQL and how to download data from a database. You'll analyze local and global temperature data and compare the temperature trends where you live to overall global temperature trends.

Project	Investigate a Dataset

In this project, you'll choose one of Udacity's curated datasets and investigate it using NumPy and pandas. You'll complete the entire data analysis process, starting by posing a question and finishing by sharing your findings.

LESSON TITLE	LEARNING OUTCOME
ANACONDA	• Learn to use Anaconda to manage packages and environments for use with Python.
JUPYTER NOTEBOOKS	• Learn to use this open-source web application to combine explanatory text, math equations, code, and visualizations in one shareable document.
DATA ANALYSIS PROCESS	• Learn about the keys steps of the data analysis process. • Investigate multiple datasets using Python and Pandas.
PANDAS AND NUMPY: CASE STUDY 1	• Perform the entire data analysis process on a dataset. • Learn more about NumPy and Pandas to wrangle, explore, analyze, and visualize data.
PANDAS AND NUMPY: CASE STUDY 2	• Perform the entire data analysis process on a dataset. • Learn more about NumPy and Pandas to wrangle, explore, analyze, and visualize data.
PROGRAMMING WORKFLOW FOR DATA ANALYSIS	• Learn about how to carry out analysis outside Jupyter notebook using IPython or the command line interface.

Course 2: Practical Statistics

Learn how to apply inferential statistics and probability to important, real-world scenarios, such as analyzing A/B tests and building supervised learning models.

Project	Analyze Experiment Results

In this project, you will be provided a dataset reflecting data collected from an experiment. You'll use statistical techniques to answer questions about the data and report your conclusions and recommendations in a report.

| LESSON TITLE	LEARNING OUTCOME
SIMPSON'S PARADOX	• Examine a case study to learn about Simpson's Paradox.
PROBABILITY	• Learn the fundamental rules of probability.
BINOMIAL DISTRIBUTION	• Learn about binomial distribution where each observation represents one of two outcomes. • Derive the probability of a binomial distribution.
CONDITIONAL PROBABILITY	• Learn about conditional probability, i.e., when events are not independent.
BAYES RULE	• Build on conditional probability principles to understand the Bayes rule. • Derive the Bayes theorem.
STANDARDIZING	• Convert distributions into the standard normal distribution using the Z-score. • Compute proportions using standardized distributions.
SAMPLING DISTRIBUTIONS AND CENTRAL LIMIT THEOREM	• Use normal distributions to compute probabilities. • Use the Z-table to look up the proportions of observations above, below, or in between values.
CONFIDENCE INTERVALS	• Estimate population parameters from sample statistics using confidence intervals.
HYPOTHESIS TESTING	• Use critical values to make decisions on whether or not a treatment has changed the value of a population parameter.

Course 2: Practical Statistics, cont.

LESSON TITLE	LEARNING OUTCOME
T-TESTS AND A/B TESTS	• Test the effect of a treatment or compare the difference in means for two groups when we have small sample sizes.
REGRESSION	• Build a linear regression model to understand the relationship between independent and dependent variables. • Use linear regression results to make a prediction.
MULTIPLE LINEAR REGRESSION	• Use multiple linear regression results to interpret coefficients for several predictors.
LOGISTIC REGRESSION	• Use logistic regression results to make a prediction about the relationship between categorical dependent variables and predictors.

Project Example > Investigate a Dataset
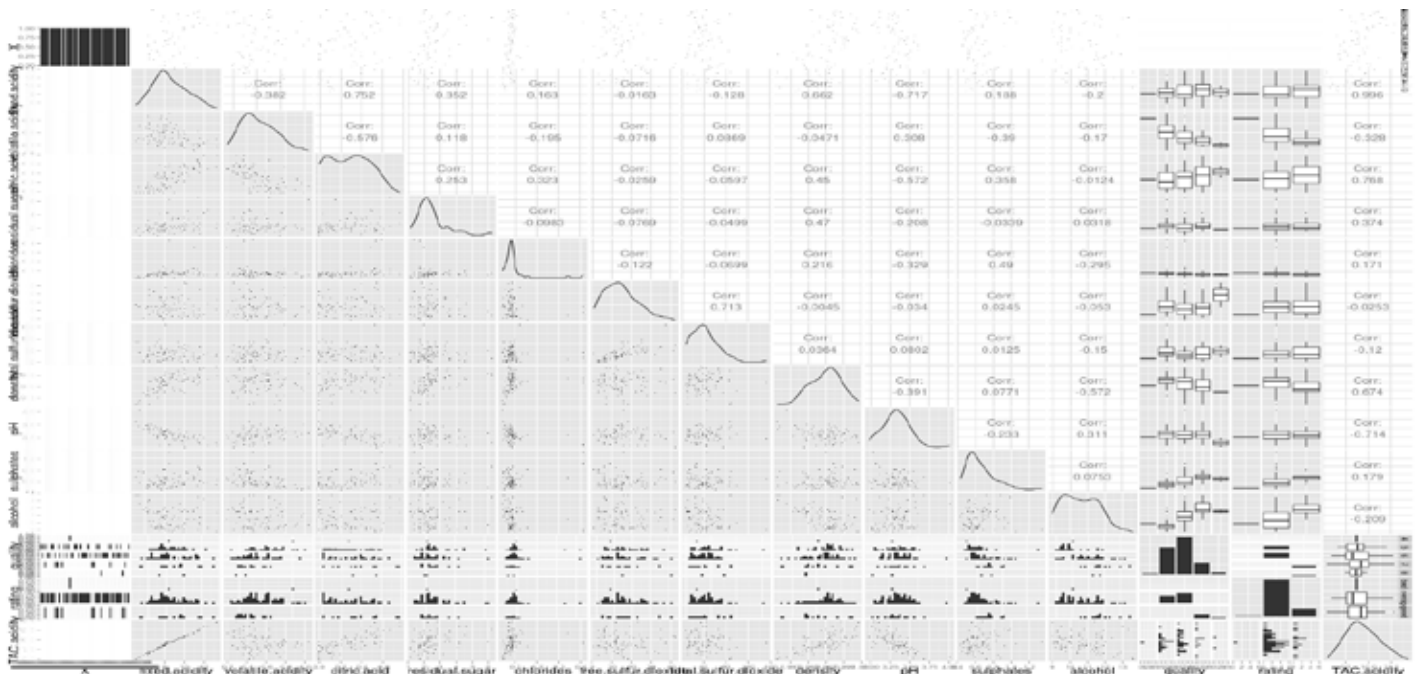
Learner-submitted Data Analyst Nanodegree course project solution

Course 3: Data Wrangling

Learn the data wrangling process of gathering, assessing, and cleaning data. Learn how to use Python to wrangle data programmatically and prepare it for deeper analysis.

Project	Wrangle and Analyze Data

Real-world data rarely comes clean. Using Python, you'll gather data from a variety of sources, assess its quality and tidiness, then clean it. You'll document your wrangling efforts in a Jupyter Notebook, plus showcase them through analyses and visualizations using Python and SQL.

LESSON TITLE	LEARNING OUTCOME
INTRO TO DATA WRANGLING	• Identify each step of the data wrangling process (gathering, assessing, and cleaning). • Wrangle a CSV file downloaded from Kaggle using fundamental gathering, assessing, and cleaning code.
GATHERING DATA	• Gather data from multiple sources, including gathering files, programmatically downloading files, web-scraping data, and accessing data from APIs. • Import data of various file formats into pandas, including flat files (e.g. TSV), HTML files, TXT files, and JSON files. • Store gathered data in a PostgreSQL database.
ASSESSING DATA	• Assess data visually and programmatically using pandas. • Distinguish between dirty data (content or "quality" issues) and messy data (structural or "tidiness" issues). • Identify data quality issues and categorize them using metrics: validity, accuracy, completeness, consistency, and uniformity.
CLEANING DATA	• Identify each step of the data cleaning process (defining, coding, and testing). • Clean data using Python and pandas. • Test cleaning code visually and programmatically using Python.

Course 4: Data Visualization with Python

Learn to apply visualization principles to the data analysis process. Explore data visually at multiple levels to find insights and create a compelling story.

Project	Communicate Data Findings

Real-world data rarely comes clean. Using Python, you'll gather data from a variety of sources, assess its quality and tidiness, then clean it. You'll document your wrangling efforts in a Jupyter Notebook, plus showcase them through analyses and visualizations using Python and SQL.

LESSON TITLE	LEARNING OUTCOME
DATA VISUALIZATION IN DATA ANALYSIS	• Understand why visualization is important in the practice of data analysis. • Know what distinguishes exploratory analysis from Explanatory analysis, and the role of data visualization in each.
DESIGN OF VISUALIZATIONS	• Interpret features in terms of level of measurement. • Know different encodings that can be used to depict data in visualizations. • Understand various pitfalls that can affect the effectiveness and truthfulness of visualizations.
UNIVARIATE EXPLORATION OF DATA	• Use bar charts to depict distributions of categorical variables. • Use histograms to depict distributions of numeric variables. • Use axis limits and different scales to change how your data is interpreted.
BIVARIATE EXPLORATION OF DATA	• Use scatter plots to depict relationships between numeric variables • Use clustered bar charts to depict relationships between categorical variables. • Use violin and bar charts to depict relationships between categorical and numeric variables. • Use faceting to create plots across different subsets of the data.
MULTIVARIATE EXPLORATION OF DATA	• Use encodings like size, shape, and color to encode values of a third variable in a visualization. • Use plot matrices to explore relationships between multiple variables at the same time. • Use feature engineering to capture relationships between variables.
EXPLANATORY VISUALIZATIONS	• Understand what it means to tell a compelling story with data. Choose the best plot type, encodings, and annotations to polish your plots. • Create a slide deck using a Jupyter Notebook to convey your findings.
VISUALIZATION CASE STUDY	• Apply your knowledge of data visualization to a dataset involving the characteristics of diamonds and their prices.

Our Nanodegree Programs Include:
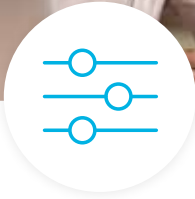
Pre-Assessments

Our in-depth workforce assessments identify your team's current level of knowledge in key areas. Results are used to generate custom learning paths designed to equip your workforce with the most applicable skill sets.

Dashboard & Progress Reports

Our interactive dashboard (enterprise management console) allows administrators to manage employee onboarding, track course progress, perform bulk enrollments and more.

Industry Validation & Reviews

Learners' progress and subject knowledge is tested and validated by industry experts and leaders from our advisory board. These in-depth reviews ensure your teams have achieved competency.

Real World Hands-on Projects

Through a series of rigorous, real-world projects, your employees learn and apply new techniques, analyze results, and produce actionable insights. Project portfolios demonstrate learners' growing proficiency and subject mastery.

Our Review Process

Real-life Reviewers for Real-life Projects

Real-world projects are at the core of our Nanodegree programs because hands-on learning is the best way to master a new skill. Receiving relevant feedback from an industry expert is a critical part of that learning process, and infinitely more useful than that from peers or automated grading systems. Udacity has a network of over 900 experienced project reviewers who provide personalized and timely feedback to help all learners succeed.

Vaibhav
UDACITY LEARNER

"I never felt overwhelmed while pursuing the Nanodegree program due to the valuable support of the reviewers, and now I am more confident in converting my ideas to reality."

now at
CODING VISIONS INFOTECH

All Learners Benefit From:

Line-by-line feedback for coding projects	Industry tips and best practices	Advice on additional resources to research	Unlimited submissions and feedback loops

How it Works
Real-world projects are integrated within the classroom experience, making for a seamless review process flow.

- Go through the lessons and work on the projects that follow
- Get help from your technical mentor, if needed
- Submit your project work
- Receive personalized feedback from the reviewer
- If the submission is not satisfactory, resubmit your project
- Continue submitting and receiving feedback from the reviewer until you successfully complete your project

About our Project Reviewers

Our expert project reviewers are evaluated against the highest standards and graded based on learners' progress. **Here's how they measure up to ensure your success.**

900+	1.8M	3	4.85 /5
Expert Project Reviewers	**Projects Reviewed**	**Hours Average Turnaround**	**Average Reviewer Rating**
Are hand-picked to provide detailed feedback on your project submissions.	Our reviewers have extensive experience in guiding learners through their course projects.	You can resubmit your project on the same day for additional feedback.	Our learners love the quality of the feedback they receive from our experienced reviewers.

UDACITY
FOR ENTERPRISE