

# Data Analyst Syllabus



---

## Contact Info

While going through the program, if you have questions about anything, you can reach us at [enterprise-support@udacity.com](mailto:enterprise-support@udacity.com). For help from Udacity Mentors and your peers visit the Udacity Classroom.

## Nanodegree Program Info

Learn to organize data, uncover patterns and insights, make predictions using machine learning, and clearly communicate critical findings.

### Prerequisite Skills

A well-prepared learner is able to:

- Proficient in object oriented programming fundamentals in Python
- Experience working with Data in Python using NumPy and Pandas
- Proficient in SQL

### Required Hardware

- Webcam
- Microphone

### Required Software

- Sublime Text 2.0.2+
- Python 3.6
- Jupyter Notebook
- Pandas 0.18.1
- NumPy 1.11
- SciPy 0.18.1
- Matplotlib 1.5.3
- Pandasql 0.7.3
- Seaborn 0.7.1
- SQLite 3.x
- RStudio 1.0+
- R 3.4
- Tableau Public 10.2
- Anaconda 4.7 or latest
- GIT BASH 2.23 or latest

- SQLAlchemy

**Version:** 4.0.0

**Length of Program:** 104 Days\*

*\* This is a self-paced program and the length is an estimation of total hours the average student may take to complete all required coursework, including lecture and project time. Actual hours may vary.*

## Part 1: Welcome to the Nanodegree program!

Welcome to the program! In this part, you'll get an orientation into using our classroom and services. You'll also get advice for making the best use of your time while enrolled in this program.

### **Project: Explore Weather Trends**

In this project, you will analyze local and global temperature data and compare the temperature trends where you live to overall global temperature trends.

## Part 2: 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: Investigate a Dataset**

Choose one of Udacity's curated datasets, perform an investigation, and share your findings.

## Part 3: 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 A/B Test Results**

You will be working to understand the results of an A/B test run by an e-commerce website. Your goal is to work through to help the company understand if they should implement the new page design.

## Part 4: 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

Gather data from a variety of sources and in a variety of formats, assess its quality and tidiness, then clean it. Showcase your wrangling efforts through analyses and visualizations.

### Supporting Lessons

Lesson	Summary
<b>Introduction to Data Wrangling</b>	Identify each step of the data wrangling process (gathering, assessing, cleaning) through a brief walkthrough of the process. The dataset for this lesson is an online job postings dataset from Kaggle.
<b>Gathering Data</b>	Gather data from various sources and a variety of file formats using Python. Rotten Tomatoes ratings, Roger Ebert reviews, and Wikipedia movie poster images make up the dataset for this lesson.
<b>Assessing Data</b>	Assess data visually and programmatically for quality and tidiness issues using pandas. The dataset for this lesson is mock Phase II clinical trial data for a new oral insulin called Auralin.
<b>Cleaning Data</b>	Using pandas, clean the quality and tidiness issues you identified in the "Assessing Data" lesson. The dataset is the same: mock Phase II clinical trial data for a new oral insulin called Auralin.

## Part 5: Data Visualization

Learn to apply sound design and data visualization principles to the data analysis process. Learn how to use analysis and visualizations to tell a story with data.

### Project: Communicate Data Findings

Choose a dataset, either your own or a Udacity-curated dataset, and perform an exploratory data analysis using Python. Then, create a presentation with explanatory plots that conveys your findings.

## Part 6: Congratulations and Next Steps



Udacity

Generated Thu May 21 03:34:31 PDT 2020