**Week 1**

**What is MLOps?**

| **Reading Title** | **Brief Description** |
| --- | --- |
| [MLOps: Continuous delivery and automation pipelines in machine learning](https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning#mlops_level_1_ml_pipeline_automation) | This article serves as a primer article to different approaches for an MLOps pipeline for a given application. It leverages the knowledge and experience of Google's history of ML services to illustrate some common pitfalls of MLOps projects. |
| [What is DevOps?](https://www.atlassian.com/devops) | This article provides an overview of the history, definition, and benefits of DevOps, as well as linking out to a number of useful resources for applying DevOps principles to your organization or workflow. |
| [Sowing the seeds of ethical AI: 4 tasks to stay on track](https://cloud.google.com/blog/transform/4-tasks-to-ensure-your-companys-ai-is-ethical) | This article outlines four guiding principles that should be considered and implemented during the development and delivery of any public-facing AI process. |
| [Build, Test and Deploy ETL solutions using AWS Glue and AWS CDK](https://aws.amazon.com/blogs/big-data/build-test-and-deploy-etl-solutions-using-aws-glue-and-aws-cdk-based-ci-cd-pipelines/) | This article provides a step-by-step process to create a continuous integration/continuous delivery solution using AWS Glue and AWS CDK. It also provides all of the example code to create this workflow in your own AWS account, and encourages you to use this process as a model to create your own AWS solutions. |

Most of the readings for this lesson focus on defining guiding principles or important processes that we'll be using throughout the rest of this course and specialization. The last article walks through a step-by-step process of creating a cloud-based CI/CD software solution that is housed in AWS. While optional, if you choose to follow these processes alongside the instructions, you'll experience workflows and pipelines that you can use to create your own solutions in the future.

## Key Concepts in MLOps

| **Reading Title** | **Brief Description** |
| --- | --- |
| [Increase reliability in data science and machine learning projects with CircleCI](https://circleci.com/blog/increase-reliability-in-data-science-and-machine-learning-projects-with-circleci/) | This article goes over the importance of automated testing, and setting up CircleCI to help accommodate that process. Additionally it also provides the starter code to get you up and running with CircleCI, and links to useful documentation about further utilization of the service. |
| [Machine Learning operations maturity model](https://learn.microsoft.com/en-us/azure/architecture/example-scenario/mlops/mlops-maturity-model?WT.mc_id=academic-0000-alfredodeza) | This article outlines the MLOps maturity model, which clarifies the important processes, technologies, and people that are present in progressively more mature MLOps environments. It's beneficial to understand what might be required to successfully implement and maintain an MLOps environment of a given maturity. |
| [DevOps tech: Continuous delivery](https://cloud.google.com/architecture/devops/devops-tech-continuous-delivery) | This article explains the organizational and technical benefits of implementing a Continuous Delivery workflow. Additionally, it illustrates some of the typical obstacles that organizations have to implementing a CD workflow, and proposes practical approaches to overcome them. |
| [Detect data drift on datasets](https://learn.microsoft.com/en-us/azure/machine-learning/v1/how-to-monitor-datasets?tabs=python&view=azureml-api-1) | This article explains the concept of data drift, and some workflows to help analyze and mitigate it. Additionally, it gives some example code and step by step instructions for utilizing some Azure Machine Learning workflows to help mitigate data drift in your own models. |
| [Automate MLOps with SageMaker Projects](https://docs.aws.amazon.com/sagemaker/latest/dg/sagemaker-projects.html) | This is a link to the Amazon Sagemaker projects page, which is a subsection of the Amazon Sagemaker developer guide that can walk you through step by step instructions on using Sagemaker to create any number of MLOps projects. More than anything, this is useful documentation to come back to when you'd like to create a Sagemaker MLOps project. |
| [Databricks Feature Store](https://learn.microsoft.com/en-us/azure/databricks/machine-learning/feature-store/) | This short article explains what exactly a feature store is, and some of the benefits of Microsoft's Databricks feature store in particular. Additionally, it links to a number of useful APIs and example notebooks that can give you some hands-on practice to illustrate the value of feature stores. |

## Foundations of End to End MLOps with Web Apps and Microservices

This Lesson's additional readings are focused on different tools that you will encounter when building web applications. Going through these additional readings will give you some experience with different CI/CD product offerings (from Google, Amazon, Github, etc.) that will be invaluable tools in creating your own pipelines and applications. Additionally, understanding CI/CD pipelines will be critical to effectively creating some of the artifacts in this course and specialization, so it's worth taking the time to briefly read this documentation.

| **Title** | **Brief Description** |
| --- | --- |
| [Build a Web Application with Flask](https://paiml.com/docs/home/books/minimal-python/chapter06-build-web-apps-flask/) | This is a chapter from a book authored by this course's instructors! It walks you through using the Flask web framework to create a web application that can use your python code. If you follow along with these instructions, you'll wind up with a functional web app. |
| [MLOps for Azure](https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/manage/mlops-machine-learning) | MLOps with Azure Machine Learning offers an efficient way to manage the entire machine learning lifecycle. By following this reading, students will learn how to create reproducible models, reusable training pipelines, and simplify model packaging, validation, and deployment. Moreover, they will gain insights into best practices for quality control, A/B testing, and automating the retraining process. This knowledge will help students improve the quality and consistency of their machine learning solutions, making them better prepared for real-world applications. |
| [Introduction to Microservices](https://www.nginx.com/blog/introduction-to-microservices/) | This article introduces the concept of Microservices, as well as the benefits and drawbacks of using a microsystem-based approach. Understanding these kinds of systems gives a lot of insight into how many large web applications function, and helps solve the problem of scalability within systems you may design. |
| [Building a Go App on App Engine](https://cloud.google.com/appengine/docs/standard/go/building-app) | This tutorial will take you through the process of making a bare-bones Go App using google's AppEngine and Google Cloud. Using this free service will give you experience in deploying an application and using AppEngine to manage your deployments. |
| [Getting Started with App Runner](https://docs.aws.amazon.com/apprunner/latest/dg/getting-started.html) | This documentation is an introduction to Amazon's App Runner, which is a system that allows you to turn a container or source code into a web service using AWS. This is just a cursory overview of using App Runner to deploy whatever application you've made, and when you're creating web applications it will be important for you to consider which delivery method or product best suits your use case. |
| [Learn Github Actions](https://docs.github.com/en/actions/learn-github-actions/understanding-github-actions) | This is the full documentation for Github actions, so you can choose to dive in as deeply as you desire. Github actions is a platform that can help you manage your continuous integration/continuous delivery pipelines, and configure workflows to trigger whenever specific events occur in a given repository, among other advanced functionality. |

## Week 2

## Doing Data Science Your First Day

| **Title** | **Brief Description** |
| --- | --- |
| [Data Science with Pandas](https://paiml.com/docs/home/books/minimal-python/chapter07-data-science-pandas/) | This textbook chapter will introduce you to the streamlined data science workflow that we use, in addition to the Pandas library, which is one of the most useful tools for doing data science with python. |
| [Data Science Libraries](https://paiml.com/docs/home/books/minimal-python/chapter08-data-science-libraries/) | This textbook chapter walks you through using some of the most common libraries that are used for data science with python. |
| [Google Earth Colab Notebook](https://developers.google.com/earth-engine/guides/python_install-colab) | This article and example Colab notebook walks you through importing the Google Earth Python API into your own Colab notebook, as well as having you do some of the steps of Exploratory Data Analysis on your own. The steps of exploring and plotting your data are integral to data science, and learning these workflows can help you in future data science applications. |
| [Tensorflow Colab Notebooks](https://cloud.google.com/tpu/docs/colabs) | This article is simply a repository for a whole group of colab notebooks that are designed to provide examples and let you interact with Google's Clout TPUs (or tensor processing units) to understand how you might leverage AI services in your own workflows. |
| [Introductory Concepts in Python, IPython and Jupyter](https://colab.research.google.com/drive/1jPhYOyvweOJVW3Engci6GcnVEGd_UhGI) | to Python, Jupyter, and Colab, covering essential concepts and techniques. The content is divided into two main parts: Part 1.1 focuses on introductory concepts such as procedural statements, data structures, strings, and arithmetic operations. It also covers Python dictionaries, lists, and sets, and how to use them effectively. Part 1.2 delves into functions, including writing functions, function arguments, closures, partial functions, generators, decorators, and lambdas. The exercise also demonstrates applying functions to Pandas DataFrames and integrating Python with other languages. Throughout the exercise, students are exposed to Colab-specific features like magic commands, mounting Google Drive, and using forms. This exercise is crucial for students as it equips them with foundational knowledge and practical skills in Python programming and working with Jupyter and Colab, which are vital for pursuing a career in machine learning, AI, or data science. |
| [Understanding Libraries, Classes, Control Structures, and Regular Expressions](https://colab.research.google.com/drive/1dmTMM-zmGToCmcaedyNj93X-dLZeg2KG#scrollTo=4J2lIggJLUMS) | This colab notebook walks through many of the basic parts of writing python, and serves as a very cursory overview of the important parts of the python language that you'll need to understand to do meaningful python projects. |
| [IO Operations in Python and Pandas](https://colab.research.google.com/drive/1sORrIFB8t9NbW9Nhko-EgFI1HWOKwn28#scrollTo=qhdbSG-xtuPU) | This colab notebook walks through using the Pandas library to do input/output operations. These are integral to data science and data engineering applications, as most people will tend to work with large bits of data in something like a Jupyter notebook, but you must understand file operations before you're able bring your dataset into that environment. |
| [AWS Cloud-Native Python for ML/AI](https://colab.research.google.com/drive/1hl0aHbJqoJIyC1P3OwK9UljP7bupTKEv#scrollTo=lujOt1rIipPQ) | This colab notebook walks you through the process of using a myriad of Amazon Web Services in collaboration with your own colab notebooks to be able to start to build out a project that leverages AI APIs and ML workflows. |
| [Additional Topics in Data Science](https://github.com/noahgift/functional_intro_to_python#additional-topics) | As part of this Coursera course, students are encouraged to browse through a curated list of Colab notebooks and repositories to gather inspiration for their own projects. These resources cover a wide range of topics, including Python programming recipes, immutability, managed machine learning systems, Internet of Things, and best practices in software development. By exploring these materials, students can gain insights into writing clean and maintainable code, testing Jupyter notebooks, implementing continuous integration with CircleCI, and ensuring reliability in data science and machine learning projects. These resources serve as a valuable starting point for students to develop their projects while adhering to industry standards and best practices, ultimately enhancing their understanding of the course material and fostering the development of real-world problem-solving skills. |

## Optimization, Heuristics and Simulations

| **Title** | **Brief Description** |
| --- | --- |
| [Greedy Optimization Repo](https://github.com/nogibjj/Coursera-MLOps-C2-lab4-greedy-optimization) | This GitHub repo features a lab called "Greedy Optimization," aimed at teaching students how to enhance command-line tools and scripts for improved performance and functionality. By engaging with this lab, students will learn to modify a greedy change tool for better robustness and explore error handling and script improvements. They will also work with a traveling salesman algorithm to optimize simulation runs. The lab provides a Python script implementing the greedy coin change algorithm, offering practical experience in working with real-world code. |
| [Linear Optimization at Scale Repo](https://github.com/noahgift/tornado_rabbitmq_linear_optimization/blob/master/README.md) | The assigned article presents an insightful introduction to linear optimization using the Pyomo library in Python. However, it is important to note that the code examples provided are outdated. Students should focus on understanding the underlying concepts and principles of linear optimization, rather than relying on the specific code implementations presented. For those who are motivated to delve deeper, it is highly recommended to consult the latest Pyomo documentation to access up-to-date information about the SDK. By doing so, students can stay informed about recent developments and best practices in utilizing the Pyomo library for solving complex optimization problems in various fields. |
| [Archived | Linear optimization in Python, Part 1: Solve complex problems in the cloud with Pyomo](https://developer.ibm.com/articles/cl-optimizepythoncloud1/) | Reading continued |
| [Operations Research Repo](https://github.com/noahgift/or) | This source code repository, Operations Research, provides a collection of operations research algorithms, with its origins tracing back to a course taught by Dr. David L. Woodruff at the UC Davis MBA program. The repository contains valuable resources, including a Traveling Salesman Problem (TSP) solution using randomized start and greedy path approaches, Colab Notebook exploring greedy algorithms and random choices, and a series of YouTube lectures on business analytics. Students are encouraged to engage with these materials and actively participate in the exercises provided.  Please note that to run the code, you'll need to utilize Docker for containerization. Detailed instructions for building and running the container are provided in the README. Additionally, you'll find examples of greedy algorithms and random choices, as well as YouTube resources on business analytics and containerization. By exploring and experimenting with the contents of this repository, students can deepen their understanding of operations research concepts and techniques, further enhancing their problem-solving skills in various domains. |
| [Introduction to Machine Learning Problem Framing](https://developers.google.com/machine-learning/problem-framing) | In this assigned reading, you'll learn how to assess whether machine learning (ML) is a suitable approach for addressing a specific problem. This course, provided by Google, covers essential concepts and techniques for outlining an ML solution. You'll gain valuable insights into the initial stages of designing and implementing ML solutions. Understanding problem framing is crucial for successful ML projects, as it ensures that the chosen approach aligns with the problem at hand and the desired outcomes. |

## Machine Learning and AI in Practice

| **Title** | **Brief Description** |
| --- | --- |
| [Review Notebooks in Core Statistics for Data science](https://github.com/noahgift/core-stats-datascience) | The "Core Statistics for Data Science" tutorial is an interactive learning resource that features a repository full of Jupyter Notebooks. These notebooks cover a wide range of data science topics, including data science workflow, supervised learning trends, clustering, GMM, PCA, recommendations, and network analysis. To engage in this exercise, students can access the code examples provided in the repo and follow along with the concepts discussed in the Notebooks. |
| [Practitioner’s Guide to MLOps Whitepaper](https://cloud.google.com/resources/mlops-whitepaper) | The "Practitioner's Guide to Machine Learning Operations (MLOps)" offers an overview of the MLOps life cycle, focusing on essential processes and capabilities for successful ML-based systems. Students will learn about continuous training pipelines, model deployment, data management, and model governance, gaining valuable insights into managing ML projects and enhancing their MLOps skills. |
| [Notebook:  Managed Machine Learning Systems and Internet of Things](https://github.com/noahgift/managed_ml_systems_and_iot/blob/master/PROD_Chips.ipynb) | This interactive tutorial takes students through some of the current issues in computer hardware related to machine learning, such as vectorized code and custom silicon. |
| [Clustering](https://developers.google.com/machine-learning/clustering) | This course on clustering provides students with valuable hands-on experience in data preparation, similarity measures, and the k-means algorithm for clustering data. By completing this course, you will not only gain a deeper understanding of clustering techniques and their applications in machine learning, but also develop practical skills for implementing and evaluating clustering results. This course is designed to complement your existing knowledge in machine learning problem framing, data preparation, and feature engineering, and will ultimately enhance your ability to develop effective ML models. |
| [Introduction to Machine Learning](https://developers.google.com/machine-learning/intro-to-ml) | Welcome to Introduction to Machine Learning, a concise 20-minute course that provides an introduction to the foundational concepts of machine learning (ML). In this course, you will learn about the different types of ML and the key concepts of supervised machine learning. You will also gain insight into how solving problems with ML differs from traditional approaches. Please note that this course does not cover the implementation of ML or working with data. |
| [Machine Learning Glossary](https://developers.google.com/machine-learning/glossary) | As a student in the Coursera course, you can frequently use the Machine Learning Glossary as a valuable supplementary resource. This glossary provides definitions for general machine learning terms as well as terms specific to TensorFlow. To enhance your understanding, you can filter the glossary by selecting a topic from the Glossary dropdown in the top navigation bar. Keep an eye out for the hatching bird icon, which signifies definitions aimed at ML newcomers. Utilize this glossary to reinforce your learning and clarify any unfamiliar terms or concepts throughout the course. |
| [Machine Learning Crash Course](https://developers.google.com/machine-learning/crash-course) | Students can use the Machine Learning Crash Course with TensorFlow APIs as a complementary resource to their main Coursera course. While you don't need to complete the entire Crash Course, it offers an excellent opportunity to randomly explore new lessons and expand your knowledge. The course provides a fast-paced, practical introduction to machine learning, featuring video lectures, real-world case studies, and hands-on practice exercises. By selectively diving into specific topics that interest you or enhance your main course material, you can enrich your learning experience and gain a broader understanding of machine learning concepts. |