



# Reading: Course 3 resources and citations

---

## Module 1: Find and share stories using data

### Resources

#### Case study: Deloitte

- [Deloitte](#)
- [Deloitte provides a range of services for client projects](#)

#### Reference guide: The EDA process

- [Institute for Ethical AI & Machine Learning](#)

### Citations

#### Case study: Deloitte

- Us, D. (2020, December 2). [Data in motion by Deloitte: Connected operations](#) [Video]. YouTube

#### Reference guide: The EDA process

- [The responsible machine learning principles](#). (n.d.). The Institute for Ethical AI & Machine Learning
- Zempekakis, I. (2022, March 30). [Regulation and ethics in data science and machine learning](#). Medium

## Module 2: Explore raw data

### Resources

#### Reference guide: Import datasets with Python

- [BigQuery Sandbox](#)
- [An overview of importing data in Python](#)
- [How to connect to BigQuery from a Colab](#)

#### Reference guide: Pandas methods for the discovery of a dataset

- [A list of Pandas dataframe functions](#)

#### Reference guide: Datetime manipulation

- [module in Python's standard library](#)
- [NumPy's datetime classes](#)
- [Pandas' datetime classes](#)
- [pandas.to\\_datetime\(\) documentation](#)
- [pandas dt accessor documentation](#)

## Reference guide: Pandas tools for structuring a dataset

- [df.merge\(\)](#)
- [pd.concat\(\)](#)
- [df.join\(\)](#)
- [df.select\\_dtypes\(\)](#)
- [df.sort\\_values\(\)](#)
- [df.iloc\[\]](#)
- [df.loc\[\]](#)
- [Pandas documentation to describe parameters in Python functions](#)
- [W3schools provides explanations for Python functions in an easy-to-understand way](#)

## Histograms

- [plt.hist\(x, bins=10, ...\)](#)
- [sns.histplot\(x, bins, binrange, binwidth ...\)](#)

## Citations

### EDA using basic data functions with Python

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]

### Date string manipulations with Python

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]

### Reference guide: Date/time manipulation

- Miyazaki, M. (2021, April 29). [Python date and time conversion cheat sheet](#). DEV Community
- Child, D. (2011, October 19). [Python Cheat Sheet](#). Cheatology
- [Python 3.10.6 documentation](#). (n.d.). Python

## Module 3: Clean your data

### Resources

#### Data deduplication with Python

- [Look at Pandas documentation to learn more about the parameters of the argument field](#)
- [W3 Schools: Pandas - removing duplicates](#)

#### Protect the people behind the data

- [Harvard Business School: 5 principles of Data Ethics for Business](#)
- [Dataversity: Data Governance and Data Quality](#)

### Reference guide: Data cleaning in Python

- [df.info\(\)](#)
- [pd.isna\(\) / pd.isnull\(\)](#)
- [pd.notna\(\) / pd.notnull\(\)](#)
- [df.fillna\(\)](#)

- [df.replace\(\)](#)
- [df.dropna\(\)](#)
- [df.describe\(\)](#)
- [sns.boxplot\(\)](#)
- [df.astype\(\)](#)
- [Series.cat.codes](#)
- [pd.get\\_dummies\(\)](#)
- [LabelEncoder\(\)](#)

## Citations

### Data deduplication with Python

- [pandas 1.4.3 documentation](#). (n.d.). Pandas

### Protect the people behind the data

- Ghosh, P. (2019, July 25). [Data governance and data science: What is the intersection?](#) DATAVERSITY.
- [Home Depot Annual Report 2021](#). (2021). Home Depot
- Southecal, P. (2020, January 24). [Data compliance: An integral part of business analytics](#). Data Science Central

### Reference guide: How to handle outliers

- Bonthu, H. (2022, July 21). [Detecting and treating outliers: Treating the odd one out!](#) Analytics Vidhya

### Identify and deal with outliers in Python

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]

### Label encoding in Python

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]

### Other approaches to data transformation

- Sethi, A. (2020, June 25). [One-hot encoding vs. label encoding using Scikit-Learn](#). Analytics Vidhya
- [When to use One Hot Encoding vs LabelEncoder vs DictVectorizer?](#) (2015, December 19). Data Science Stack Exchange.

### Input validation with Python

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]

# Module 4: Data visualizations and presentations

## Resources

### Tableau Public overview

- [set up your Tableau Public profile](#)
- [download](#)
- [save and share your work](#)
- [Tableau.com](#)
- [Tableau Help page](#)
- [Set up data sources](#)
- [Design charts and analyze data](#)
- [Welcome to Tableau Public](#)
- [A step-by-step guide to get you started on your own data viz journey](#)
- [A step-by-step guide for analyzing data and publishing a viz to your Tableau Public profile](#)

### How to sign on to Tableau Public

- [login credentials](#)
- [Tableau Public FAQ](#)
- [reset password](#)
- [Tableau Public home page](#)
- [Welcome to Tableau Public](#)
- [Resources page](#)
- [Technical specifications and storage requirements](#)
- [The Tableau Public Discover page](#)
- [The Tableau Public Blog](#)

### Download your datasets and begin presenting with Tableau

- [Tableau Public](#)

### Follow-along guide: Work with Tableau, Part 1

- [set up your Tableau Public profile](#)

### Follow-along guide: Work with Tableau, Part 2

- <https://public.tableau.com/s/>
- [set up your Tableau Public profile](#)

### Follow-along guide: Craft compelling stories with Tableau

- <https://public.tableau.com/s/>
- [set up your Tableau Public profile](#)

### The top five data visualization resources

- [Tableau Public](#)
- [Resources](#)
- [Viz of the Day](#)

- [PowerBI](#)
- [guided learning](#)
- [online workshops](#)
- [PowerBI data visualization community](#)
- [Information is Beautiful](#)
- [online workshops](#)
- [Storytelling with Data](#)
- [community](#)
- [Python Graph Gallery](#)
- [Yan Holtz](#)
- [Dataviz Inspiration](#)

### **Follow-along guide: Present like a pro with Tableau**

- <https://public.tableau.com/s/>
- [set up your Tableau Public profile](#)

### **Activity: Design a bar graph that tells a story in Tableau Public**

- [Tableau Public](#)

### **Activity: Build an interactive dashboard in Tableau Public**

- [Tableau Public](#)

## **Citations**

### **Tableau Public overview**

- [Discover](#). (n.d.). Tableau Public
- Google Data Analytics Professional Certificate: Meet Tableau. [Video]. (n.d.). Google
- [Tableau Help](#). (n.d.). Tableau

### **The visualization life cycle**

- [Get the data](#). (n.d.). Inside Airbnb

### **Work with Tableau, Part 1**

- [Tableau Public](#). (n.d.). Tableau Public

### **How to sign on to Tableau Public**

- [Tableau Public](#). (n.d.). Tableau Public

### **Work with Tableau, Part 2**

- [Build Data Views for Accessibility](#). (n.d.). Tableau
- [Tableau Public](#). (n.d.). Tableau Public

### **The top five data visualization resources**

- [Data visualization: Microsoft Power BI](#). (n.d.). Microsoft Power Bi

- Holtz, Y. (n.d.). [Python graph gallery](#). The Python Graph Gallery
- [Information is beautiful](#). (n.d.). Information is Beautiful
- [Tableau: Business intelligence and analytics software](#). (n.d.). Tableau
- [Storytelling with data](#). (n.d.). storytelling with data

### **Craft compelling stories with Tableau**

- NOAA. (n.d.). [Cloud-to-ground lightning strikes](#) [Dataset]
- [Tableau Public](#). (n.d.). Tableau Public

## **Module 5: Course 3 end-of-course project**

### **Resources**

#### **Activity: Create your Course 3 Automotidata project**

- [2017\\_Yellow\\_Taxi\\_Trip\\_Data.csv](#)
- [Tableau Public](#)

#### **Activity: Create your Course 3 TikTok project**

- [Tableau Public](#)

### **Citations**

#### **Course 3 end-of-course project overview**

- [About TLC](#). (n.d.). NYC Taxi & Limousine Commission
  - [FHV trip record data](#). (n.d.). NYC Taxi & Limousine Commission
-