🚀 Pre-Course Setup & Resources Guide

Welcome to **Python for Basketball Data Science!**

This guide ensures you have all the necessary tools and accounts set up before you begin Week 1. This course uses **Google Colab**, which means you don't need to install any software on your personal computer—you just need a Google account.

---

1. ⚙️ Technical Setup: Google Colab

Google Colaboratory (Colab) is a free, cloud-based platform that allows you to write and execute Python code through your web browser.

Task 1: Create a Google Account

- **Requirement:** You must have a **Google account (Gmail)** to access and save your work in Google Colab and Google Drive.
- **Action:** If you do not have one, create a new Gmail account.

Task 2: Access Google Colab

- **Action:** Go to the Google Colab Website.
- **Recommendation:** Bookmark this link!
- **File Storage:** All your notebooks (.ipynb files) and saved data (.csv files) will be automatically stored in a folder called **Colab Notebooks** within your **Google Drive**.

Task 3: Test a Notebook

1. Open Google Colab.
2. Click **"New notebook"** in the bottom right corner of the welcome dialog.
3. In the first code cell, type: `print("Setup Complete!")`
4. Click the **"Play" button** (▶️) next to the cell.
5. If you see the output `Setup Complete!`, you are ready to start Week 1.

---

2. 📘 Essential Python/Pandas Terminology

Throughout the course, we will use specific terms to describe data and code actions. Familiarize yourself with these core concepts:

Term	What It Is	Course Context
Variable	A name used to store a specific piece of data (e.g., a player's points or a calculated rate).	**Week 1**
Function	A named block of reusable code that performs a specific task (e.g., calculating effective_fg).	**Week 2**
DataFrame	The primary data structure in the Pandas library. It's a two-dimensional table, like a spreadsheet.	**Week 3**
Series	A single column within a DataFrame.	**Week 3**
Imputation	The technique of filling in missing data points (NaN) with calculated estimates (like the mean or median).	**Week 7**
Feature Engineering	The process of creating **new metrics** (or features) from raw data (e.g., turning PTS, FGA, FTA into TS%).	**Week 9**
Correlation (r)	A statistical measure of how closely two variables relate. Closer to 1.0 means they move strongly together (positive relationship).	**Week 6**

---

3. 🧩 Using the Weekly Notebooks

Each week, you will perform the required tasks directly within a Jupyter Notebook.

How to Start Each Week

1. Open your Google Drive.
2. Navigate to the **Colab Notebooks** folder.
3. **Find the new assignment notebook** (e.g., Week_X_Assignment.ipynb).
4. **Crucial Step: Always save a copy before starting!** Go to **File > Save a copy in Drive**. Rename the new copy (e.g., Week_X_Assignment_MyName). This prevents you from losing the original blank version.

Key Commands (Shortcuts)

Action	Menu	Shortcut
Run Current Cell	Runtime	Shift + Enter
Add New Code Cell	Insert	Ctrl/Cmd + B
Add New Text Cell	Insert	Ctrl/Cmd + M

---

4. 📚 Data Science Stack (The Libraries)

We will use the following Python libraries for virtually all the work in this course:

Library	Standard Alias	Purpose	Weeks Used
Pandas	pd	Data cleaning, manipulation, filtering, and grouping.	**All Weeks (3-10)**
NumPy	np	High-performance numerical and statistical calculations (mean, correlation).	**All Weeks (6-10)**
Matplotlib	plt	Base library for generating plots and charts.	**All Weeks (5, 6, 8, 9, 10)**
Seaborn	sns	High-level library for creating complex, attractive statistical visualizations (like heatmaps).	**All Weeks (5, 8, 9, 10)**
Scikit-learn	sklearn	Machine Learning tasks (scaling, splitting data, training models).	**Week 10**

You are now fully prepared for the technical and conceptual journey of the course! Good luck!