EXPLORE AI

2409 PTDS | Regression Project Kick-Off

EXPLORE AI ACADEMY

Introduction

Project Objectives

Educational Goal	Introduction to practical applications of Regression Analysis in Python focusing on forecasting average temperature in Agriculture
Project Goal	Kick-Start the Team project. Apply learned concepts and technical skills to a live dataset.
Collaborative Goal	Build on the use of collaborative tools introduced Python and expand on those tools to include: Trello for project management.
Final Outcome	Ability to deliver comprehensive analysis and presentation of findings, demonstrating technical and competence and team work.

The Project:

- The aim is to analyse and predict average temperature from the agri-food sector, using data from the FAO and IPCC, to understand climate impacts and develop sustainable strategies for stakeholders including policymakers and agricultural businesses.

The Dataset:

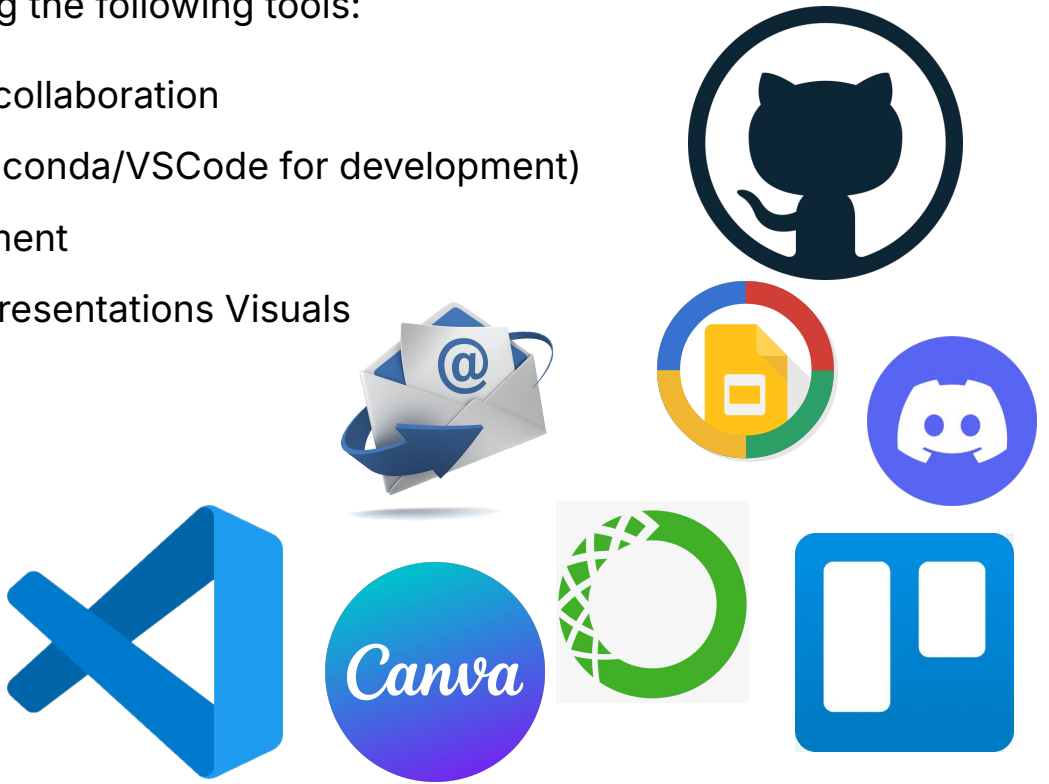
- Multiple features (30+)
- ~ 7000 records

EXPLORE AI ACADEMY

Overview of Tools

In this project we would be using the following tools:

- Github for versioning and collaboration
- Project Environments (Anaconda/VSCode for development)
- Trello for Project Management
- Google Slides/Canva for Presentations Visuals



GitHub

- GitHub & Git will be used for the project versioning and documentation.
- One team member should fork the **Project** repository and add the other team members as collaborators. **Please click here for the Project repository.**
- Add you facilitator as a collaborator. Your facilitators GitHub User names will be shared in your Project Teams Emails.
- Every team member needs to clone the repo, create a branch, and work from there. **We will be accessing contributions (commits) to the repository from each member.**
- You will have to deal with merge conflicts 🙃
- Your GitHub repo will need a ReadMe file.
  Here is a helpful link to shortcuts in Markdown.

Python Overview

- You will need to include all the packages used in a **requirements.txt** file in your GitHub repository
- In the **ReadMe you need to add instructions on how to recreate the environment using Anaconda.**
- Here are a few helpful links to get you started:
  - Managing environments
  - From conda create requirements.txt for pip3
- Exporting your conda environment:

```
conda activate <env>
conda install pip
#get list of packages and pipe to txt file
pip list --format=freeze > requirements.txt
```

Team Roles and Collaboration Strategies

Trello

- All team members will be **required** to create a free Trello account. Click here to sign up if you do not have one.
- A team member should be assigned Project Manager to manage the Trello board
- The Project Manager will be responsible for creating a new board and inviting other team members to join.
- They will also be responsible for creating list, cards, assigning tasks, using labels and checklist to keep track of of the project.
- You can learn more about how to use Trello to manage this project by watching this video.

Slide Deck

- These are tools for creating beautiful engaging slides.

Project Timelines - Continued

Regression _ Week 3

- Understand the Dataset.
- Elect a team lead, project manager (manages trello)
- Set up all other collaborative and development tools required for the project.
- Clean your dataset & Perform Exploratory Data Analysis.
- Begin working on the slide deck.

Regression _ Week 4

- Perform train-test splits.
- Train with a **minimum of 3 models.**
- Explore and decide on Model Evaluation Metrics.
- Compare and contrast models performance.
- Add all relevant information about the model to the slide deck (Model comparisons, evaluations, and others).

Project Timelines - Continued

Regression _ Week 5

- Continuation of Week 4 Deliverables

Regression _ Week 6

- Complete Deliverables and submit project files

Submit by Deadline:

Monday, 02 June 2025

EXPLORE AI ACADEMY

Key Takeaways

Deliverables:

- Trello for Project Task assignments and the entire project management.
- GitHub for collaboration and version control.
- Google Slide/Canva for presentation deck.
- Jupyter Notebook which should be:
    - Well formatted to contain all relevant sections
    - A robust and well annotated EDA.
    - At least 3 regression models.
    - A model comparison, selection, and recommendation section.

Important links

Please find below, important links:

- Project Repository: Click here
- Facilitator Github Usernames: Click here
- Managing Environments: Click here
- Creating environments from **requirements.txt** using "**conda create**": Click here
- Jupyter notebook markdown cheatsheet: Click here
- Sign-up to Trello: Click here
- Video on how to set-up your Trello board: click here