NU Faculty Web Scraper

This project scrapes faculty data from NU Lahore Campus (http://lhr.nu.edu.pk/faculty/) using Python. It extracts faculty details like names, designations, HEC approval status, department, email addresses, and profile images from various departments and exports the results to structured CSV files.

# 📑 Table of Contents

* ✨ Features
* 📦 Requirements
* ⚙️ Setup Instructions
* 🚀 How to Run
* 📁 Output
* 📂 Project Structure
* 📜 License

# ✨ Features

- Scrapes all departments from the NU faculty page dynamically.  
- Extracts:  
 - Faculty Name  
 - Designation  
 - HEC Approval status  
 - Department  
 - Email address  
 - Profile image URL  
 - Faculty ID  
- Exports data to individual CSV files for each department.

# 📦 Requirements

Python 3.7+  
Jupyter Notebook (optional)  
Libraries:

* - requests  
  - beautifulsoup4  
  - pandas

Install all dependencies using:

***pip install -r requirements.txt***

# ⚙️ Setup Instructions

1. Clone this repository:

***git clone https://github.com/yourusername/nu-faculty-scraper.git  
cd nu-faculty-scraper***

2. Install required Python packages:

***pip install requests beautifulsoup4 pandas***

# 🚀 How to Run

Run the Jupyter Notebook using:

***jupyter notebook WebScarping.ipynb***

Or convert to a script and run:

***jupyter nbconvert --to script WebScarping.ipynb  
python WebScarping.py***

# 📁 Output

The notebook will create CSV files for each department, e.g.:

* fsc.csv  
  ee.csv  
  cv.csv  
  ss.csv  
  fsm.csv

Each CSV file contains structured faculty data with fields like:

* - ID  
  - Name  
  - Designation  
  - HEC Approval (True/False)  
  - Department  
  - Email  
  - Image URL

# 📂 Project Structure

nu-faculty-scraper/  
├── WebScarping.ipynb # Main notebook  
├── \*.csv # Output files for each department  
├── requirements.txt # List of dependencies  
└── README.md # You're here!

# 📜 License

This project is licensed under the MIT License.

👨‍💻 Built with ❤️ by M. Azib Naeem