Ordine dello sviluppo del progetto:

1. **Web Scraping Capology.ipynb** (per il web scraping dei dati dal sito Capology: si ottengono i dati nella cartella “dati ancora separati/Capology”)
2. **FBREF e Transfermarkt.R** (per l’import dei dati dai siti FBREF e Transfermarkt: si ottengono i dati nella cartella “dati ancora separati/Fbref”, e i dati nella cartella “dati ancora separati/Transfermarkt”; il file Rdata FBREF e Transfermarkt.RDATA contiene l’environment una volta eseguito lo script di R)
3. **Data Integration e data quality.ipynb** (per integrazione dei dati, esportazione nei dati definitivi poi caricati su MongoDB, e data quality: : si ottengono i dati nella cartella “dati integrati e puliti”)

Parti di codice non utilizzate.R e Parti di codice non utilizzate.ipynb sono due files in cui sono contenute altre funzioni che abbiamo utilizzato in diverse fasi del progetto, senza che poi siano state inserite direttamente nella stesura finale del progetto.