StatLock v2.2 — CFB Roster Ingestion (Ourlads) — Data Dictionary

Created 2025-10-06

Purpose: Standardize NCAA FBS roster ingestion from Ourlads for modeling, props, and player usage analytics.

Version	2.2
Module	CFB_Rosters_Ourlads
Upstream	https://www.ourlads.com/ncaa-football-depth-charts/
Inputs	FBS_master_roster_urls.csv; Ourlads team roster pages
Outputs	cfb_rosters_ourlads_fbs.csv; cfb_rosters_ourlads_fbs.jsonl

Fields

Name	Type	Required	Example	Description
team_name	string	Yes	Ohio State Buckeyes	Official team name as rendered on Ourlads ro
conference	string	Yes	Big Ten	Conference of the program at time of scrape.
stadium	string	No	Ohio Stadium (102,780)	Stadium name and capacity when listed on O
surface	string	No	Grass / Shaw Sports Turf	Playing surface as listed on team roster page
head_coach	string	No	Ryan Day	Head coach as listed under team info.
off_scheme	string	No	Spread / Pro Spread / Air Raid	Offensive scheme string from Ourlads.
def_scheme	string	No	4-2-5 / 3-3-5	Defensive scheme string from Ourlads.
updated_at	datetime|string	No	2025-09-30 13:30 ET	Timestamp text shown as 'Updated' on the ro
player_name	string	Yes	Cade Klubnik	Player full name.
jersey	string|int	No	2	Jersey number; keep as string to preserve lea
pos	string	Yes	QB	Roster position abbreviation.
height	string	No	6020	Height as displayed; often in scouts-style inch
weight	string|int	No	210	Weight in pounds.
class	string	No	JR / RS SO	Eligibility class reported by Ourlads.
high_school	string	No	Westlake	High school name where listed.
hometown	string	No	Austin, TX	Hometown as listed.
source_url	string	Yes	https://www.ourlads.com/ncaa-foo Roster depth-charts/roster.aspx?id=91533&s=o	Roster page URL used for parsing.

Primary Keys: team_name, player_name, pos, source_url

Uniqueness Notes: Duplicate player names may appear across seasons/programs; include source_url in PK for uniqueness.

Quality Checks:

• Drop rows with missing player_name or pos.
• Normalize whitespace; trim commas and extra spaces.
• If jersey contains non-numeric characters, keep as-is (string).
• Keep height as raw string; a separate transform can convert to inches.

Transform Notes:

• Create derived fields later: height_in, bmi (requires height parse), class_bucket (FR/SO/JR/SR).
• Conference realignment is fluid; store as of scrape date in a separate table if needed.

License/Use: Use Ourlads data for academic research; respect site terms and robots directives.

Contact: Maintainer: StatLock v2 (for Zynda PhD project)