

Dataset Nutrition Label

Finding Correlation Between Athletes Salary and Difficulty

- About
 - Our final project's goal is to try to find the leading cause of what allows a lot of athletes to earn such a high paying salary. Is it because of the skill, toughness, or popularity of the sport? Or is it because of another unknown factor? Furthermore we hope that our dataset also creates a platform that allows regular people to see how much athletes truly make and what jobs they do for it.
- Data Created
 - 2020
- Data collected
 - 1990 - 2020
- Created by
 - Toughest Sports Dataset - Rishi Damarla
 - Forbes Highest Paid Athletes 1990 - 2020 - Parul Pandey
 - Final Project - Nate Valdez, Arjan Sethi, and Justin Hing
- Links
 - <https://www.kaggle.com/datasets/rishidamarla/toughest-sports-by-skill/>
 - <https://www.kaggle.com/datasets/parulpandey/forbes-highest-paid-athletes-1990-2019>
- Content
 - 2 Datasets
- Background
 - Our project represents sports and athletes from the whole world.
 - Our data is collected from some of the highest paid athletes across multiple different sports from 1990 to 2020.
 - In our toughest sports dataset the acronyms mean:
 - END - Endurance, STR - Strength, PWR - Power, SPD - Speed, AGI - Agility, FLX - Flexibility, NER - Nerve, DUR - Durability, HAN - Hand Eye Coordination, and ANA - Analytical Aptitude
 - Data from the Toughest Sports Dataset was created by sports scientists from the US Olympic Committee that averaged and totalled a degree of difficulty for each sport on a scale from 1 - 100.
- Issues
 - The Forbes dataset had issues with the format of how the sports in the sports column were labeled. The sports would alternate from being capitalized or uncapitalized. They also didn't match up to the Toughness dataset's sports column when it came to sports like auto racing, baseball, MMA, and various others.
- Collections
 - Both data from the datasets came from ESPN or Topendsports.com