

Estimating the Hubble Constant Using SDSS DR16 Galaxy Data

Project Overview

This project applies real astronomical data from Sloan Digital Sky Survey (SDSS-IV Data Release 16) to estimate the Hubble constant (H_0) — a key parameter of the Λ CDM (Lambda-Cold-Dark-Matter) cosmological model.

Using redshift measurements for thousands of galaxies, the analysis quantifies the linear relationship between galaxy velocity and distance, validating the universe's ongoing expansion.

Objective

Estimate the Hubble constant (H_0) using observational data and compare it to the standard Λ CDM predictions ($\approx 67\text{--}74 \text{ km/s/Mpc}$).

Data Source

- Dataset: SDSS-IV Data Release 16 (DR16) – Spectroscopic Galaxy Sample
- Accessed via: <https://skyserver.sdss.org/dr16/en/tools/search/sql.aspx>
- Key fields used: z (Redshift), ra & dec (Coordinates), $modelMag_r/g/i$ (Apparent magnitudes)

Methods & Tools

1. **Data Wrangling:** Cleaned DR16 CSV export and filtered valid redshifts ($z > 0$).
2. **Derived Features:** Computed velocity ($v = z \times c$) and distance ($d = v/H_0$).
3. **Regression Modeling:** Fitted linear regression to estimate H_0 .
4. **Visualization:** Scatter and regression plots of Hubble's Law.
5. **Validation:** Compared estimated H_0 with Λ CDM predictions.

Results

- Estimated $H_0 \approx 69 \text{ km/s/Mpc}$ (within Λ CDM range 67–74)
- Model $R^2 = 0.82$ (strong linear correlation)
- Observation: Distant galaxies show higher recession velocity, confirming cosmic expansion.

Insights

1. Estimated H_0 supports Λ CDM cosmology.
2. Outliers may result from local gravitational effects or noise.
3. Strong R^2 confirms Hubble's Law validity in DR16 data.
4. Analysis reinforces the universe's large-scale linear expansion.

Tech Stack

Python, pandas, numpy, matplotlib, seaborn, scikit-learn, Jupyter Notebook

Resume Summary Line

Hubble Constant Estimation using SDSS DR16 Data — Analyzed 1,000+ galaxy observations using Python (pandas, scikit-learn) to estimate $H_0 \approx 69 \text{ km/s/Mpc}$. Validated Λ CDM cosmology predictions through regression and visualization.

Next Steps

- Explore higher- z galaxies for cosmic acceleration trends.
- Extend model using non-linear Λ CDM curvature terms.
- Integrate larger SDSS datasets for deeper regression precision.