

## MLB 球员薪资预测

在职业体育领域，球员的薪资不仅关系到俱乐部的财务管理，也影响到球队的整体战斗力和人才储备。尤其是在美国职业棒球大联盟（Major League Baseball, MLB）中，球员薪资的制定是一项复杂而关键的决策过程。俱乐部在签约球员或决定续约时，通常会参考球员的历史表现、职业生涯统计数据、场上贡献、所处的位置、球队的预算情况等多种因素。球员也会根据自己的表现寻求市场上与其能力相匹配的合同金额。

本数据集（data.csv）源于著名的《Introduction to Statistical Learning》教材，包含了 1986 年 MLB 球员的一系列表现数据及其年薪信息。我们的目标是使用线性回归（stepwise regression 或 LASSO）的方法建模，并探讨不同变量与薪资之间的关系。

### 数据说明：

数据集包含以下变量：

- Salary：球员年薪（单位：千美元），**为响应变量 y**。
- AtBat：本赛季击球次数。
- Hits：本赛季击中球的次数。
- HmRun：本赛季本垒打次数。
- Runs：本赛季得分。
- RBI：本赛季打点数（即击球使队友得分的次数）。
- Walks：保送（四坏球上垒）次数。
- Years：球员在 MLB 的职业年数。
- CAtBat：职业总击球次数。
- CHits：职业总击中数。
- CHmRun：职业总本垒打。
- CRuns：职业总得分。
- CRBI：职业总打点。
- CWalks：职业总保送。
- PutOuts：野手防守中完成的接杀出局数。

- Assists: 防守助攻次数。
- Errors: 防守失误次数。

### **作业要求:**

- 1, 本作业具有开放性, 但必须使用完整且严谨的统计方法进行分析。
- 2, 该报告要求具有一定的可读性, 即使用普通阅读者可以理解的方式撰写报告。
- 3, 学生可以自行选择使用 PYTHON 或者 R, 代码应包含在报告里。
- 4, 作业要求**必须**上传 PDF (如果你使用的软件无法直接保存为 PDF, 请用截屏工具将其转化为图片, 之后插入 word, 再转成 PDF) 。
- 5, **报告必须在 6 页以内**, 超过 6 页将会被扣分。请合理进行编排, 格式字体字号不限。
- 6, 该报告需要独立完成, 但可以使用大模型以及互联网等工具。