蔡炟婧

+86 17305098017 | cdajing@163.com



网页CV

教育背景

对外经济贸易大学（保研） **应用统计 | 硕士** **2026.09 - 2028.07**

天津财经大学 **经济统计学 | 本科** **2022.09 - 2026.07**

专业技能

- 取数与数据处理：精通SQL（多表关联、窗口函数）| 熟练使用Excel（VLOOKUP）| 熟练使用Python（Pandas、NumPy）
- 数据仓库基础：熟悉Hive（数仓构建）| 熟悉Hadoop（生态基础）
- 业务与用户分析：掌握问卷分析、联合分析、用户分群与RFM | 熟悉AARRR、漏斗分析、异动与归因分析
- 统计验证与建模：熟悉相关性分析、假设检验与A/B测试 | 掌握回归、分类、聚类、时序分析与模型调优
- 可视化与呈现：掌握Tableau（看板搭建）、Matplotlib（绘图）与PPT（报告呈现）

实习经历

凯度Kantar（市场调研咨询领域Top） **定量调研组 | 数据分析实习生** **2026.03–2026.07**

日常工作

- **排查问卷路径错误** 整理竞品与品牌活动资料，核对题目、选项及翻译，按不同受访者路径测试筛选、跳转和显示逻辑。记录问题交编程同事修改，随后复测，并参与回收中期检查及开放题补码。
- **对齐分群与报告口径** 为比较不同客群，按分析问题补充分群条件，手动取数或协调数据处理团队，核对分母与比例。将差异整理成图表及初步结论，按审核意见补充行业资料。

实习项目

- **识别品牌形象变化 | 传统浏览器** 用 4A 框架跟踪主品牌及小说、网盘、短剧赛道，对照上期与竞品表现。在品牌属性维度用 BIP 处理知名度与样本基数差异，梳理目标客群中的显著变化与相对优势。
- **挖掘采购痛点 | AI 电商** 面向 B2B 采购场景，从问卷回收表提取创业探索者与已有经营者的挑战提及比例，分群排序并与总体对照，呈现两类人群的优先痛点，支持传播活动分析。
- **统一取数与质检口径 | 跨境支付** 在跨境旅行支付研究启动时，将图表对应到题号、筛选条件、分群和分母。参考往期项目及本期要求参与配额调整，制定开放题质检规则，覆盖模板化回答及表述与评分矛盾等情形。

项目经历

瑞幸咖啡联名产品用户消费意愿探究 **项目负责人** **2023.10–2024.05**

- **提炼联名偏好指标** 主导研究和抽样设计，从访谈提炼 **5 类 18 项指标**，用正交设计筛选代表性产品组合。回收 1,277 份问卷，保留 **1,148 份有效样本（89.9%）**。
- **识别优先触达客群** 用 K-prototypes 聚类和 RFM 价值分析区分三类消费者，定位占样本 **34.1%** 的重要发展客群，提炼差异化的营销触达重点。
- **推导联名产品组合** 用联合分析比较属性偏好，以 IPA 确定改进优先级。提出 **25–30 元（不含 30 元）的民族风澳白**、艺术家或设计师联名及包装类赠品组合，并给出分群营销建议。获**市调国赛一等奖、天津市总冠军**。

蚂蚁森林用户评论情感分析 **项目负责人** **2024.03–2024.06**

- **建立评论的功能分类** 为了解个人碳账户用户在记录低碳行为时遇到的功能问题，独立用 Python 清洗、分词并去重，保留 **29,319 条**贴吧评论。用 NVivo 三级编码，将 13 个概念归入行为量化、平台功能和社交感受三类。
- **定位功能体验问题** 用 LSTM 识别情感，并按词典把评论匹配到功能。绿色能量相关评论的负向比例为 **35.75%**，高于正向的 16.51%。据此提出解释能量计算规则、完善互动规范等建议。

基于 Kaggle 洪水数据的洪水概率预测模型 **项目负责人** **2024.07–2024.08**

- **验证五项指标的预测效果** 针对部分场景难以收齐 20 项指标的问题，对比汇总统计量和随机森林筛选两种五指标方案，离线 MAE 分别为 **0.0374**、**0.0369**，据此保留全指标方案。
- **评估全指标预测模型** 独立完成清洗和训练集构建，建立 8 种回归基线并训练前馈神经网络。全指标首轮离线 **R^2 0.8637**、**MAE 0.0148**，为 74.5 万条样本生成洪水概率。

荣誉奖项

本研累计获 28 项荣誉（国家级 5 项、省部级 10 项），涵盖数学建模、统计建模、市场调查与大数据分析。

- 2024 年华数杯研究生数学建模竞赛（第 1 作者） 国家级二等奖
- 2026“华为杯”第二十三届中国研究生数学建模竞赛（第 3 作者） 国家级二等奖
- 2024 年高教社杯全国大学生数学建模竞赛（第 1 作者） 天津赛区一等奖
- 第十八届“挑战杯”天津市大学生课外学术科技作品竞赛（第 1 作者） 天津赛区二等奖
- 第二届全国大学生大数据分析技术技能大赛（个人赛） 天津赛区二等奖