高级统计学期末大作业要求

2024-11-27

**一、大作业目的**

本次大作业旨在全面检验学生对高级统计学所学知识的理解与应用能力，同时培养团队协作意识和解决实际问题的能力。通过完成大作业，学生将系统性地将课堂所学知识与现实问题相结合，从而提升分析、处理数据的能力。

**二、大作业组成**

大作业要求同学4人作为一组，自选题目来完成。

**三、大作业选题要求**

为确保选题的多样性和实践性，大作业选题需遵循以下几点要求：

1. **数据来源**：
   * **工厂或企业合作：**联系当地企业或实习单位，获取生产、物流或质量管理相关数据。
   * **公开数据集：**使用国际知名平台提供的工业相关数据集，如：
     1. Kaggle：如生产优化、供应链模拟数据集。
     2. UCI机器学习库：如设备故障数据、生产工艺数据。
     3. 政府统计局或行业组织：如制造业报告、物流行业分析数据。。
2. **选题方向**：
   * 鼓励选择具有实际意义且可以通过统计分析得出有价值结论的课题。**要求和工业工程专业或未来研究相关，例如**：
     1. **工人疲劳与绩效关系**：分析工人任务持续时间、间隔时间与生产效率之间的关系，探讨合理的工时安排。
     2. **生产数据预测分析**：利用工厂大数据（如传感器数据），进行设备状态预测、产品需求预测等。
     3. **设备故障预测**：基于历史维护数据，使用时间序列或回归分析预测设备故障发生的概率，优化维护策略。
     4. **供应链需求预测：**基于历史需求数据，供应商数据，库存数据等，预测未来的客户需求，供应链风险、价格波动等。
   * 选题应具备一定的统计分析深度，例如包含对变量的探索性分析、假设检验、统计建模等内容。
3. **数据处理与分析要求**：
   * **数据分析**：进行探索性数据分析（EDA），绘制不同图像来表现数据分布性质，对数据进行描述性统计分析。

统计建模：应用统计学方法对数据进行深入建模与分析，结合选题需求选择适当的建模方法（如线性回归、时间序列分析等），包括但不限于本课程讲授的统计建模与分析方法。

* + **结果解释**：展示分析结果，并需要结合实际问题背景进行科学解释，得出有意义的结论。
  + **工具使用**：分析过程中需使用统计分析软件（如R、Python、Matlab等），并在报告中提供关键代码或操作说明。

1. **创新性**：
   * 鼓励学生在选题和分析方法上体现一定的创新性，例如开发新指标、对现有方法进行改进，或者结合多个数据源进行综合分析。

**四、大作业PPT及报告要求**

1）每组需提交PPT（上传至网络学堂，截止时间为15周周日（12月22日）晚23:59分）：**PPT需包括背景介绍、方法分析（包括EDA，分析方法介绍，分析方法改进等）、阶段性成果、结论(包括方法的优点、缺点、未来报告里要改进地方及加入的内容)**。

2）每组提交报告一份（上传至网络学堂，截止时间为18周周日（2025年1月12日）晚23:59分）：报告中英文均可，无字数限制。报告需包括完整的选题背景、**方法分析（包括EDA，分析方法具体介绍等），**分析结果、结论与建议等内容，并**描述清楚每个小组成员的工作内容**。

**五、课堂展示时间**

每组在第十六周上课进行小组展示，每组展示时长为8分钟。具体小组展示顺序后续会公布。

