

# 最终评估：小型项目

开始时间：2023年10月16日下午2点

---

## 对小型项目的最终评估的说明

基于课程材料，本评估旨在评估你的理解能力、批判性思维能力和分析能力。在你开始之前，请仔细阅读下面的说明。

### 一般说明：

#### 项目工期：

- a. 全日制学生：7天。
- b. 兼职学生：14天

请明智地管理你的时间来完成这个项目。

允许的材料：你可以在网上参考教科书、课堂笔记资源，以及任何其他你认为必要的学习材料。

答案格式：请输入您的答案清晰和清晰。如果您包含了图表或图表，请确保它们是整齐的和适当的标签。我们将不接受手写的答案。

### 问题分解：

#### 问题1：解决计算机升级后的问题

##### 政府使用规范性分析法

完成任务并回答与案例1相关的任何问题。

#### 问题2：用数据和可视化来讲故事

说明：只选择下面两个案例研究中的一个，并回答与所选案例相关的所有子问题。

**案例研究方案2A：我们应该投资哪些房产？**

或

**案例研究方案2B：马来西亚的军事开支**

#### 问题3：开发一个机器学习模型来实现

##### 预测疾病类型

完成任务并回答与案例1相关的任何问题。

### 要记住的重要要点:

剽窃: 任何任务中的任何剽窃都将立即被取消资格。您的最终评估将不被接受, 并给予0%的分数)。你不允许从人工智能生成器或其他学生的作品中复制答案。

### 避免剽窃:

1. 引用你的资料来源: 如果你使用外部资源, 参考资料, 或引用在你的项目中, 确保正确地引用它们。剽窃包括复制别人的作品和使用他们的想法而不给予适当的赞扬。
2. 使用你自己的语言: 在解释概念或想法时, 用你自己的语言来表达它们。如果你从别人的作品中大量借鉴, 你可以解释和引用资料来源。
3. 原创性是关键: 你的项目应该展示你最初的原创想法、研究和分析。剽窃会破坏你的学术成长和你的工作的信誉。

### 协作:

1. 协作讨论: 请随意与同行讨论概念、分享资源和交换想法。协作学习可以丰富你的理解力和创造力。
2. 尊重的交流: 尊重彼此的观点和意见。不同的观点往往会导致创新的解决方案和方法。
3. 独立工作: 在鼓励讨论的同时, 你的工作应该反映出你独特的理解和努力。避免从对等点上复制解决方案、代码或内容。

请记住, 这个项目的目标不仅仅是为了产生一个结果, 而是为了提高你的技能、知识和批判性思维。拥抱合作和学习的机会, 但要确保你的贡献是诚实和真实的。祝你的项目好运。我相信你有能力创造出真正非凡的东西。

引用: 如果你指的是特定的来源、书籍或文章, 请在您的答案中提供适当的引用。

清晰: 确保你的答案是结构良好的, 清楚地表达你的想法。为了清晰, 可以根据需要使用标题、要点和段落。

复习: 在提交你的考试之前, 花几分钟时间复习一下你的答案的准确性、完整性和连贯性。

### 提交说明:

提交截止日期：

-全日制学生：2023年10月23日上午8点。

-兼职student：8.00上午，2023年10月8日。

逾期提交将不被接受。

提交格式：扫描或转换您的答案为一个PDF文档。请以以下格式命名您的文件：

“矩阵号。然后将所有处理过的答案、数据集和相关文件压缩为

“YourFirstName\_Matric\_Number”。zip”。

谢谢您的合作。祝你的迷你项目好运！

## 问题1：使用规范分析来解决政府的升级计算机问题

马来西亚行政单位（MAU）是马来西亚著名的政府机构之一，负责实现公共部门的“现代化和改革”。你是在支持主任决定升级旧电脑。主任同意，为了降低成本，在使用旧的显示器和机壳时，只为每个政府机构的电脑升级主板（打印电路板（PCB））和cpu，比购买新的电脑更便宜。

采购过程将以成本采购流程为基础。这种类型的过程用于MAU想要购买固定价格的东西。总监依靠你来决定毛需要花多少钱来购买项目，然后使用这个成本数字作为指导对项目进行投标。

在月向潜在竞标者发布招标通知后，月收到了两家大型电子制造商的报价。第一家公司被称为Samsul Co.。从韩国。这家韩国公司提供一份为期两年的合同，每股董事会支付200美元，最多支付2000pcb。Samsul为2000年以上的部分提供每块板40美元的折扣，因此以160美元的价格出售，保修一年。该保证是由Samsul公司作的书面承诺。如果他们在一年内交付的PCB出现故障，他们将免费维修或更换。然而，每台更换的PCB将在一年后每台花费200美元（i. e. 保修期满）。

毛还收到了来自日本公司的第二家制造商的报价。Toshibun提供较低的价格为120美元，但要求一次性支付12万美元作为不可退还的研发和维护费，保修两年。东信承诺，如果两年内交付的PCB出现故障，他们将免费维修或更换。

MAU的工程师已经确定，MAU可以在任何混合物中使用一个品牌或两个制造商的多氯联苯的混合物，而没有兼容性问题。然而，每台电脑的组装成本是250美元。

过度工作的硬件，很少或没有停机时间，有烟雾或火灾损坏，可能会破坏电脑。没有人知道下一年有多少台电脑需要维护。但是，CIO已经收集了关于需要维护的计算机百分比的数据，如表1.0所示。

MAU的财务副总监通知您，该PCB模型将在2025年之后停产。因此，任何一次性支付给通发的费用必须基于明年的升级要求。他要求你帮助他评估具体的经济问题，以帮助他为董事做决定的建议。

表1. 0每500台计算机中PCB制造商损坏计算机的百分比 (%)

年	托施本	Samsul公
1990	19.733	17.333
1991	18.183	18.133
1992	19.900	19.933
1993	17.217	18.950
1994	19.117	18.183
1995	16.800	19.300
1996	21.133	20.767
1997	20.500	20.200
1998	20.550	20.133
1999	21.083	21.333
2000	20.267	20.733
2001	20.517	20.233
2002	21.550	20.467
2003	20.950	21.933
2004	21.667	20.167
2005	21.267	20.783
2006	21.983	22.883
2007	22.900	23.250
2008	22.267	22.633
2009	23.117	22.617
2010	23.300	22.683
2011	23.867	24.967
2012	24.417	25.083
2013	24.050	25.383
2014	23.767	25.117
2015	23.567	25.600
2016	24.383	25.417
2017	24.933	25.383
2018	23.783	26.317
2019	24.567	25.917
2020	24.250	25.917
2021	24.783	25.700
2022	23.750	25.467

**说明：**

1. 开发一个电子表格，显示升级6000台电脑的成本  
1250个多氯联苯，其余由东石提供。将您的模型保存在工作表1上的位置。  
(10分)
2. 为了最大限度地节省政府开支，找到一个解决方案，找到最佳的数量购买多氯联苯，同时仍然履行两年的合同。假设需要升级的计算机总数为6000台。你建议毛从Samsul购买多少pcb，从东石购买多少，以减少升级和维护电脑两年的成本？在工作表2上保存分析模型和预测模型。  
(40分)
3. 在审查Samsul公司时。合同，Samsul有限公司。要求增加一个条款，如果MAU购买少于1000个pcb，MAU必须支付100,000美元作为和解协议。如果MAU使用的Samsul的PCB少于1000个，MAU主任就会要求您报告经济影响。显示对工作表3的影响。  
(50分)

## 迷你项目2：用数据和可视化来讲故事

请只为这个项目选择一个任务，或任务2A或任务2B，以完成迷你项目2。

### 案例研究选项2A：我们应该在伦敦投资哪些房产？

你们的政府关联公司（GLC）一直在努力从政府获得其外国投资基金，以支持其通过其投资部门哈塔纳Luar公司（哈塔纳）增加收入的目标。

你负责准备并向哈大拿州的董事会提交你关于投资伦敦房产的分析。给定的数据集是primarilycentredaroundthehousingmarketofLondon. 它包含了许多额外的相关数据：

- 每月平均房价
- 每年一次的房屋数
- 每年售出的房屋数量
- 每年回收利用的家庭百分比
- 沙痂每年的生活满意度
- 该地区居民的年薪中位数
- 当地居民的年平均工资
- 每月发生的犯罪次数
- 每年的工作数量
- 每年居住在该地区的人数
- 面积大小，单位：公顷

你必须研究英国的地理位置，因为这些数据被伦敦那些被称为行政区的地区所分割（有一面旗帜来识别这些地区）。尽管如此，一些变量还有其他地理区域供参考（如英格兰、东北等）。

准备演示幻灯片，回答以下管理层的业务问题：

**在伦敦的哪个地区有盈利的机会。**

**应该投资吗？**

您的演示文稿必须包括一个可视化的内容(i。e. 图表)来支持您对数据的解释。选择正确的有价值图表，可以显示/可视化因素、指标或行为模式之间的联系或关系。评论任何趋势的外观和强度，这将有助于证明你的选择。您的演示文稿不得超过20张幻灯片。

(50分)

## 案例研究方案2B：马来西亚的军费开支

对军费支出和GDP数据集进行探索性数据分析，并撰写报告。在你的报告中，包括对马来西亚军费支出趋势数据的见解。解释马来西亚的军费支出与GDP之间的关系，并将其与其他东盟国家的军事支出，如新加坡、印度尼西亚和泰国进行比较。您的报告必须包括一个可视化的内容(i。e. 图表)来支持您对数据的解释。选择正确的有价值的图表，可以显示/可视化数据和行为模式之间的联系或关系。评论任何趋势的外观和强度，或将任何模式与历史事件联系起来。您的报告在回答关键的业务问题时不得超过500字，如：发生了什么？在东盟国家中，谁在军事开支上的花费最多？与东盟国家相比，马来西亚的军事支出增长了多少？GDP增长和军费开支之间有什么关系？

(50分)



### 迷你项目3：开发一个机器学习模型来预测疾病类型

祝贺你开始了开发一个机器学习模型来预测疾病类型的最终项目。这项任务不仅能刺激智力，而且有可能对医疗保健产生重大影响。要指导您完成这个过程，这里有一些逐步的说明来帮助您完成这个小项目：

祝贺你开始了开发一个机器学习模型来预测疾病类型的最终项目。这项任务不仅能刺激智力，而且有可能对医疗保健产生重大影响。要指导您完成这个过程，这里有一些逐步的说明来帮助您完成这个小项目：

#### 1. 了解任务：

- a. 为此项目定义问题语句。
- b. 你所预测的疾病是什么？
- c. 定义你的目标。

(10分)

#### 2. 数据预处理

- a. 数据清理：数据中是否缺少值、异常值和不一致？
- b. 您执行了特性选择了吗？请证明你的回答是正确的。
- c. 您是否执行了数据转换？例如，规范化、标准化数据，或将分类变量转换为数值等价物。请证明你的回答是正确的。

(10分)

#### 3. 探索性数据分析（EDA）：

- a. 可视化数据：使用图表和图表来理解变量和疾病类型之间的关系。
- b. 描述性分析：进行描述性统计分析，以从数据中获得见解。写下三个你可以从EDA中找到的见解。

(20分)

#### 4. 模型选择

- a. 选择算法：选择三种机器学习算法进行分类任务。

- b. 模型拆分：报告您所选择的模型拆分，例如70：30。80：20或60：40。将数据集分为训练集和测试集，用于模型训练和评估。

(20分)

5. 模型培训：

- a. 训练模型：使用训练数据集来训练选定的机器学习模型。
- b. 超参数调优：优化模型参数，以达到最佳性能。

(20分)

6. 模型评估：

- a. 指标选择：根据问题的性质，报告您所选择的评估指标（准确性、精度、召回率、f1分数）。
- b. 评估模型：使用测试数据集来评估模型。比较和报告他们的性能，并选择性能最好的模型。

(20分)