

成功回答

2023-12-21 18:46

计算贷款总额





2023-12-21 18:46

Show work

根据运行结果，个人贷款的总金额为3271258.0德意志马克(DM)。

2023-12-21 18:48

计算贷款的总人数



 复制



2023-12-21 18:48

Show work

根据您提供的执行结果，个人贷款的总数为1000。这意味着在`personal_loans`表中记录了1000条个人贷款数据。

计算男性和女性贷款人的数量，并计算两者的比例



2023-12-21 18:50



Show work

执行成功后，我们得到的结果显示男性借款人数量为690，女性借款人数量为310，男女借款人之间的比例为大约2.23 : 1。

这意味着在这个样本中，男性借款人的数量是女性借款人的2.23倍。这个比例可以为贷款产品和市场营销策略提供性别相关的见解。

2023-12-21 18:53

用饼状图表示男性和女性贷款人的数量和百分比



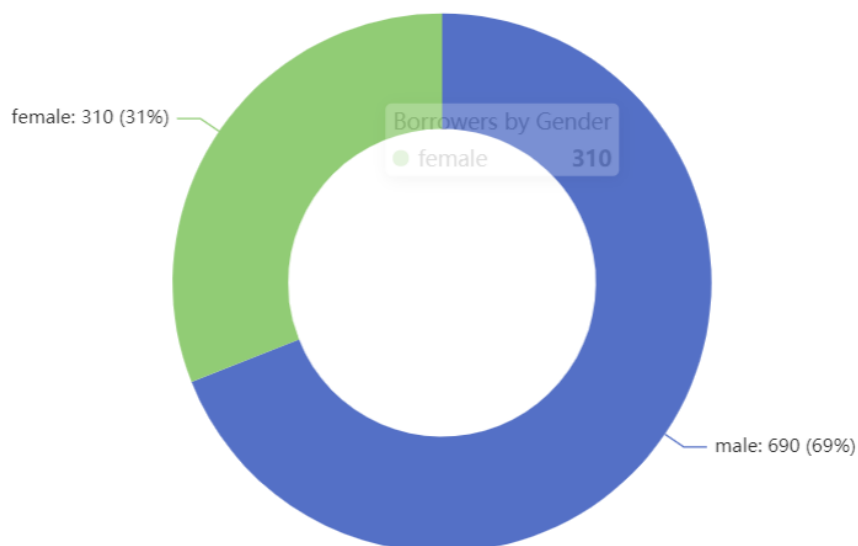
2023-12-21 18:54



Show work

Borrowers by Gender

male female



执行代码成功，并成功生成了所需的饼状图，并以 JSON 格式输出了图表信息。

饼状图展示了贷款人性别分布情况，其中男性占 690 人，女性占 310 人。比例和数量直接通过饼图进行了可视化表达。

2023-12-22 14:09

平均年龄是多少?



2023-12-22 14:10



Show work

平均年龄是35.55岁。如果您有其他问题或需要进一步的数据分析，请告诉我。

复制

2023-12-21 19:00

按年龄范围分组，计算每个年龄段的贷款人数量



2023-12-21 19:02

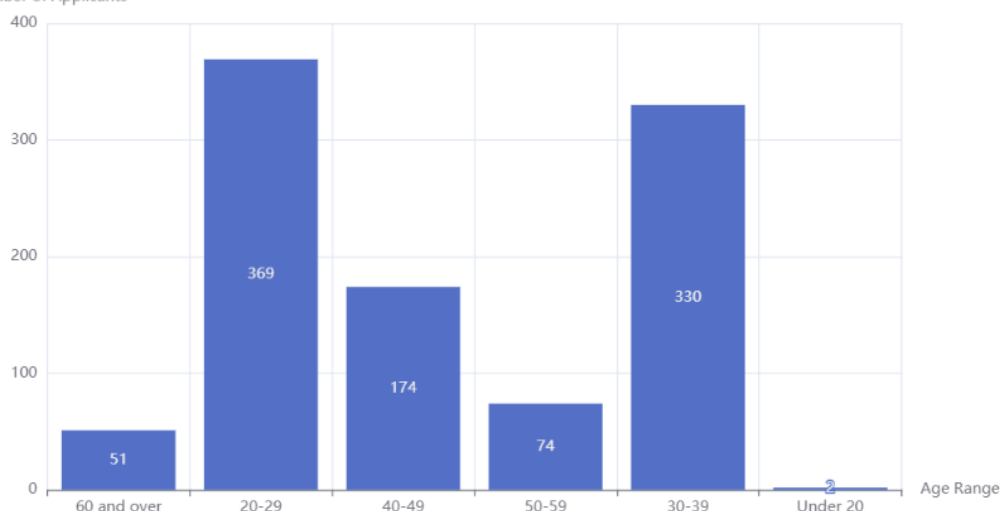


Show work

Loan Applicants by Age Range

Applicants

Number of Applicants



根据生成的图表中的数据内容，我们可以得到各个年龄段申请个人贷款的人数分布情况。具体如下：

- "20-29岁"年龄段的申请人数最多，共有369名申请人；
- 紧接着是"30-39岁"年龄段，有330名申请人；
- "40-49岁"年龄段有174名申请人；
- "50-59岁"年龄段的申请人数减少到74名；
- "60岁及以上"的群体中，有51名申请人；
- "20岁以下"的年龄段申请人数最少，仅有2名申请人。

从数据中可以看出，二三十岁的年轻人更倾向于申请个人贷款，可能是因为这个年龄段的人面临一些重要的生活事件，如购车、购房或是投资教育等，因此需要经济上的支持。随着年龄的增长，申请贷款的人数逐渐减少。年龄在20岁以下和60岁以上的人群中，申请个人贷款的人数相对较少。

这是对已生成图表中数据内容的简单介绍，更深入的分析可能需要更多关于数据、市场状况和贷款申请人的背景信息。



2023-12-22 11:07



Show work

根据执行的代码输出结果，我们可以看到不同年龄段的贷款申请数量。在提供的数据中，年龄从19岁到75岁不等，贷款申请的数量也随之变化。

通过对贷款申请数量进行分析，我们可以看到：

- 27岁的贷款申请最多，共有51次。
- 接下来是26和28岁，分别有50和43次申请。
- 从数据中还可以看出，20至30岁的年龄段申请贷款的数量普遍较多。
- 超过30岁后，申请数量逐渐减少，并随着年龄的增长，贷款申请数量明显下降。

根据这些结果，我们可以得出结论：在这批数据中，大约在20至30岁之间的年龄段最倾向于申请贷款。当然，这是基于当前数据集的分析，在实际情况中可能会受到更多因素的影响，如经济环境、贷款政策等。

这些结果为金融机构制定针对不同年龄群体的市场策略提供了依据，而年轻人可能是贷款产品的主要目标市场。金融机构可以考虑为这个年龄段的人群提供更有吸引力的贷款条件和优惠政策，以满足他们的财务需求。

2023-12-21 18:56

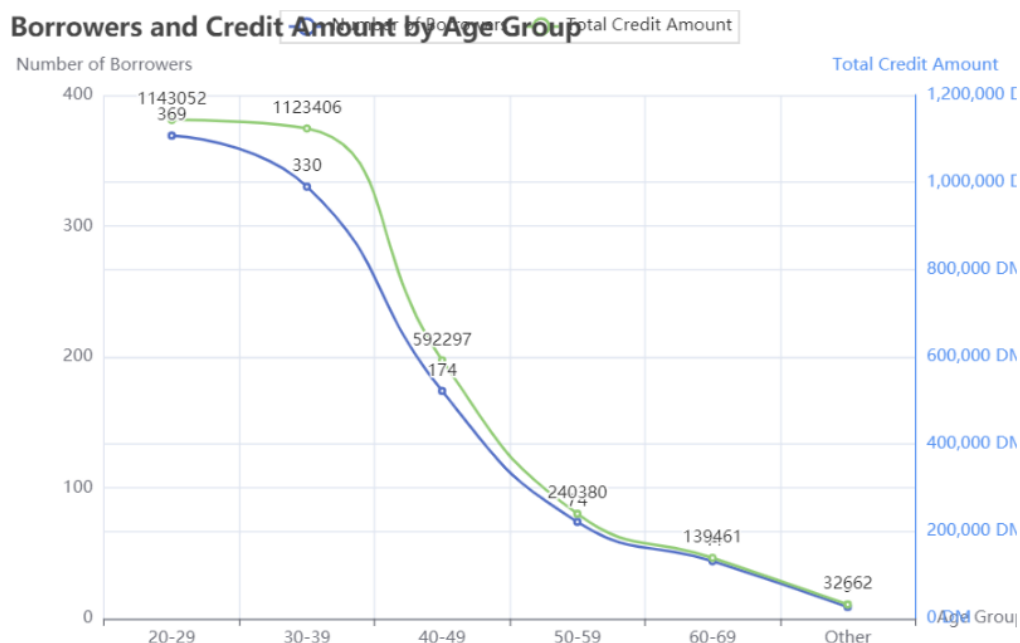
按年龄范围分组（比如20-30岁、31-40岁等），计算每个年龄段的贷款人数量和信贷额度，并用折线图表示



2023-12-21 18:59



Show work



生成的线性图表展现了不同年龄组的借款人数量和他们的信贷额度总和。该图表按年龄分组，每个组包括一个十年的范围，比如20-29岁、30-39岁，依此类推。图表的 X 轴表示年龄组，左侧 Y 轴代表借款人的数量，右侧 Y 轴则显示了该年龄组的信贷额度总和（以德国马克为单位）。

从图表中我们可以观察到，20-29岁的年龄组中借款人数量最多同时拥有最高的信贷额度总和。随着年龄的增长，借款人数量和信贷额度总和都呈现下降的趋势。比如，在60-69岁的年龄段，这两者相较于20-29岁年龄组明显减少。此外，还有一小部分数据被归类为“其他”，它包含了所有未能明确划分到具体年龄组的数据。

这样的分析结果有助于金融机构了解不同年龄段借款情况，进而制定更针对性的贷款产品策略。

2023-12-21 19:07

计算良好信用和不良信用的贷款人数量，并计算两者的百分比



复制

2023-12-21 19:08



Show work

根据查询结果，有700位贷款者具有好信用，占总贷款者的70%。同样地，有300位贷款者具有坏信用，占总贷款者的30%。这表明大部分贷款者的信用状况是好的。

2023-12-21 19:09

计算有不同储蓄和支票账户类型的贷款人数量，分别计算各类账户的百分比



2023-12-21 19:11



Show work

根据执行代码的输出结果，我们可以得到以下信息：

储蓄账户类型及借款人的数量和百分比：

- 无账户信息 (NA) : 183人, 占18.3%
- 少量 (little) : 603人, 占60.3%
- 相当富有 (quite rich) : 63人, 占6.3%
- 富有 (rich) : 48人, 占4.8%
- 适度 (moderate) : 103人, 占10.3%

支票账户类型及借款人的数量和百分比：

- 少量 (little) : 274人, 占27.4%
- 适度 (moderate) : 269人, 占26.9%
- 无账户信息 (NA) : 394人, 占39.4%
- 富有 (rich) : 63人, 占6.3%

由此可以看出，在储蓄账户类别中，“少量”占比最高，超过一半的借款人拥有少量储蓄。在支票账户类别中，“无账户信息”的比例最高，而且与“少量”和“适度”的占比差异不大。这可能暗示了支票账户分布较为均匀且有相当比例的借款人未提供账户信息。这些数据可以帮助金融机构更好地了解客户的财务状况，从而做出更加合理的贷款决策。

2023-12-21 19:21

计算不良贷款人群中不同住房类型的百分比



2023-12-21 19:23



Show work

根据上述对话中的分析，我们得知在坏账贷款人群中，不同的住房类型占比如下：

- 拥有自己住房的人占坏账总人数的62%。
- 无房产（或免费住房）的人占约14.67%。
- 租房的人占约23.33%。

可以看出，在坏账贷款人群中，拥有自己住房的人所占比例最高。

2023-12-22 11:07

不同贷款目的的申请频率如何

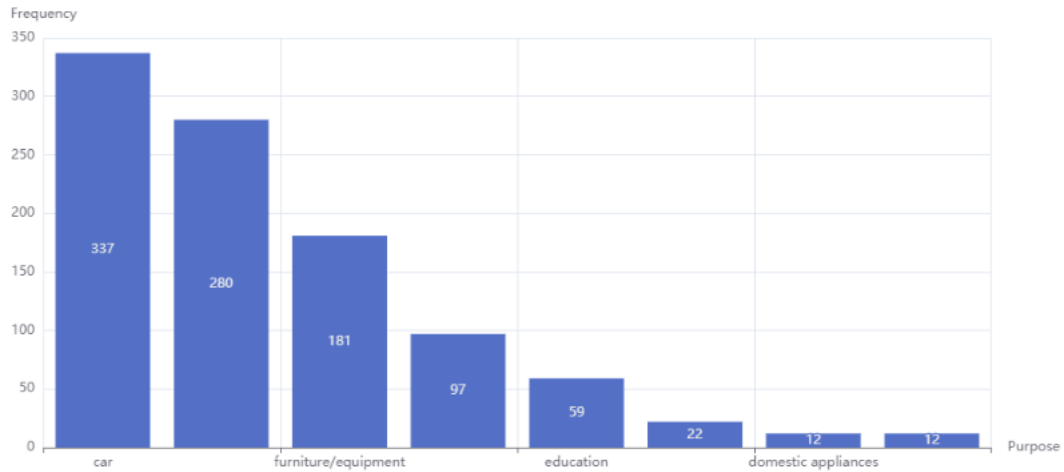


2023-12-22 11:08

Show work

Frequency of Loan Applications by Purpose

Frequency



该生成的条形图展示了不同目的的个人贷款申请频率。根据图表中的数据显示，“car”（汽车）贷款申请的频率最高，这表明汽车贷款是最受欢迎的贷款类型。其次是“radio/TV”（收音机/电视），再其次是“furniture/equipment”（家具/设备）。该图表也展示了其他贷款目的，如“business”（商业），“education”（教育），“repairs”（维修），“domestic appliances”（家用电器），以及“vacation/others”（度假/其他）。通过这些数据，我们可以了解哪些贷款目的较为常见，从而对市场需求和消费者行为进行分析。

2023-12-21 19:25

分析客户储蓄账户与是否本地居民的相关性



2023-12-21 19:27

Show work

执行结果表明，在卡方检验中得到的卡方值为4.4281，对应的P值为0.2188，自由度为3。因为P值大于0.05的常见显著性水平，我们不能拒绝原假设，即客户储蓄账户状态与是否为本地居民在统计上没有显著的关联性。

总结来说，根据现有数据和所进行的卡方检验，我们得出结论：客户储蓄账户与是否为本地居民不存在显著相关性。

2023-12-22 14:15

平均贷款金额是多少？



2023-12-22 14:16

Show work

平均贷款金额是 3271.258 德国马克 (DM)。

分析是否本地居民客户的贷款用途偏好是否存在差异



2023-12-21 19:29



Show work

分析结果显示，本地居民客户与非本地居民客户在各个贷款用途上的选择频次百分比。从给出的数据表中，我们可以大致看出两组人群在贷款用途选择上的偏好存在一定的差异。

例如，本地居民客户中选择“汽车（car）”用途的贷款相对较多，而非本地居民客户中这个比例略低，反而在“汽车（car）”用途的选择上比例较高。同样地，“业务（business）”这个选项在本地居民中的选择比例相对较低。

请注意，为了进一步验证这些差异是否具有统计学意义，可能需要进行假设检验例如卡方检验，但这可能超出了当前问题的范围，因此我们的分析只是基于现有数据。

总的来说，基于我们从数据库提取并分析的数据，本地居民客户和非本地居民客户的贷款用途偏好确实存在一些差异。具体的偏好差异可以通过以上数据表中的频次百分比进行比较理解。

是否本地居民客户的信用风险等级有何差异



2023-12-21 21:16



Show work

根据查询结果，我们可以看到非熟练非居民（Job为0）和非熟练居民（Job为1）的信用风险等级分布如下：

- 非熟练非居民（Job为0）：好信用风险有15人，坏信用风险有7人。
- 非熟练居民（Job为1）：好信用风险有144人，坏信用风险有56人。

通过计算各自的坏信用风险比例，我们可以比较两组客户的信用风险：

非熟练非居民的坏信用风险比例为 $\frac{7}{7 + 15} \approx 31.8\%$

非熟练居民的坏信用风险比例为 $\frac{56}{56 + 144} \approx 28.0\%$

观察这些数据，我们可以发现非熟练居民中信用风险被评为坏的比例稍微低于非熟练非居民。这表明本地居民（Job为1）相较于非本地居民（Job为0）有轻微的信用风险水平优势。

然而，这只是一个非常简单的分析，结论的确切性可能需要进一步的统计检验，比如卡方检验，来确认两个群体间的差异是否具有统计学上的显著性。同时，还应考虑其他可能影响信用风险评估的变量，如年龄、信用额度、贷款期限等进行更加全面的多变量分析。

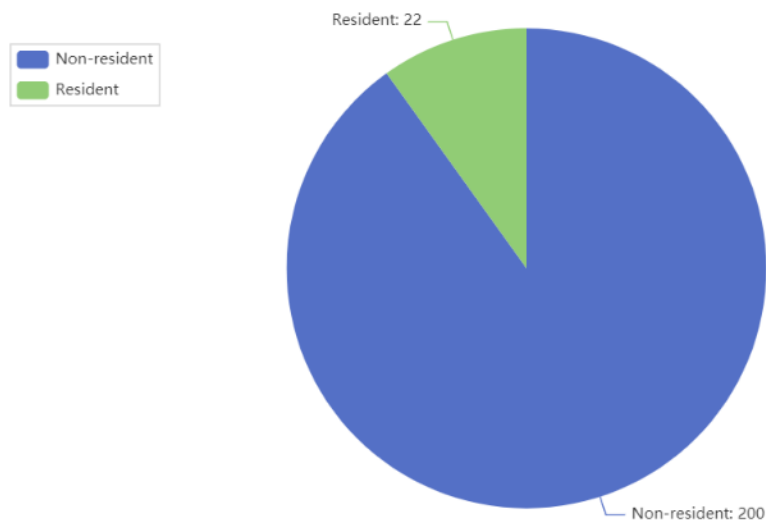


2023-12-21 21:37



Show work

Distribution of Loan Applicants by Local Resident Status



根据已生成的图表数据，我们可以得知以下信息：

- 该数据集中贷款申请者的绝大多数是非本地居民，占比达到大约90%。
- 只有一小部分的申请者是本地居民，大约占10%。

由此可以分析，非本地居民在贷款申请中占据主导地位。这个信息对于制定市场营销策略或者针对主要群体调整服务都很有帮助。此外，本地居民和非本地居民的比例也可能与特定的金融行为相关，可能需要在风险评估政策中予以考虑。

图表可以使用Pyecharts或其他可以解释所提供JSON数据结构的工具来渲染显示。

2023-12-21 21:41



复制

2023-12-21 21:41



Show work

通过查看查询结果，我们可以看到不同工作状态类别的平均信贷额度如下：

- 职业状态 3 (高技能) 的平均信贷额度最高，为 5435.49 德国马克 (DM)
- 职业状态 2 (熟练) 的平均信贷额度为 3070.97 德国马克 (DM)
- 职业状态 0 (非熟练且非居民) 的平均信贷额度为 2745.14 德国马克 (DM)
- 职业状态 1 (非熟练且居民) 的平均信贷额度最低，为 2358.52 德国马克 (DM)

根据这些数据，我们可以得出结论，职业状态为“高技能”（编号为3）的个体更有可能申请较高金额的信贷。这可能是由于他们的收入水平较高，从而可以承担更高的信贷风险和还款能力。



2023-12-22 14:34



Show work

基于查询结果，我们可以看出：

- 职业状态2（熟练工）申请贷款次数最多，为630次。
- 职业状态1（非熟练工且居住者）为200次。
- 职业状态3（高度熟练工）为148次。
- 职业状态0（非熟练工且非居住者）为22次。

因此，我们可以得出结论，‘熟练工’类别的人更倾向于更频繁地申请贷款。



2023-12-21 21:44



Show work

根据执行结果：

- ‘Checking_account’（卡方值为123.72，P值约为 $1.22e-26$ ）非常显著，表明此字段与信用风险等级有很强的关系。
- ‘Saving_accounts’（卡方值为36.10，P值约为 $2.76e-07$ ）同样表明它与信用风险等级有显著的关系。
- ‘Housing’（卡方值为18.20，P值约为0.00011）与信用风险等级也有显著的关系。

对于数值字段：

- ‘Duration’的相关系数为-0.2149，表明信用期限的增加可能与更高的风险等级相关。
- ‘Credit_amount’的相关系数为-0.1547，表明信用金额的增加可能与更高的风险等级相关。
- ‘Age’的相关系数为0.0911，影响较小，但仍然说明年龄可能与信用风险等级有一定的相关性。

其他字段如‘Sex’和‘Purpose’虽然也显示出一定的统计显著性，但它们的卡方值较小，说明与风险等级的关系可能不像上述字段那样强烈。‘Job’字段的P值为0.5966，表明与信用风险等级的关系不具统计显著性。

综上所述，主要风险因素包括：账户的当前情况（‘Checking_account’和‘Saving_accounts’）、住房状况（‘Housing’）、贷款期限（‘Duration’）和贷款金额（‘Credit_amount’）。这些因素可以在信用风险评估中得到特别关注。



2023-12-21 21:48



Show work

根据提供的数据执行结果，我们可以得出以下客户群体特征分析：

1. 在19-30岁年龄组中，男性中的“技术员”（Job值为2）人数最多，达到141人，其次为“居住的非技术员”（Job值为1）的41人。同年龄组的女性中，“技术员”也是最多的，有123人，“居住的非技术员”为其次，有28人。
2. 随着年龄组的升高，可以观察到男性“技术员”的比例依然占据主导，但在31-40岁年龄组中，“高技术专业”（Job值为3）的人数开始显著增加，达到56人。
3. 在41-50岁及51-60岁年龄组别内，女性中“技术员”和“居住的非技术员”的人数几乎持平，但男性中技术员依旧是主要工作类型。
4. 在61-70岁以及70岁以上年龄组，数量很少，可能是由于退休人群增加，但男性技术员的数量仍旧较为突出。

从上面的数据可以看出，不同年龄组的性别和职业有明显的分布特点。整体来讲，技术员（Job值为2）是最多的工作类型，在每个年龄组与性别分类中都普遍存在较大的比例。各个年龄组都显示出男性客户居多，特别是在年轻群体中。此外，年龄越大，高技术专业（Job值为3）的人数比例逐渐增加。可以认为，客户群体中，年轻男性技术员占据了较大的比例。针对这个群体的营销策略可能需要考虑这些特点。

以上分析可以为制定针对性的市场策略提供依据，例如针对技术员的工作性质和消费习惯展开营销活动。

2023-12-21 21:52

不同职业群体更倾向于选择何种类型的贷款



2023-12-21 21:54



Show work

从输出结果可以看出，不同工作状态的群体在选择贷款类型上存在一定的偏好。以下是分析报告：

- 工作状态0（非技术和非居民）：以选择购车贷款为最多，其次是商业贷款和修理贷款。
- 工作状态1（非技术和居民）：也是以购车贷款为主，其次是无线电/电视贷款和商业贷款。
- 工作状态2（技术工）：同样以购车贷款最为流行，紧随其后的是无线电/电视贷款和商业贷款。
- 工作状态3（高技能）：倾向于选择购车贷款和无线电/电视贷款，同时在选择度假或其他类贷款上也相对较多。

综上所述，购车贷款在所有工作状态下都是最受欢迎的贷款目的。而商业贷款和无线电/电视贷款在工作状态1和2中也比较常见。令人注意的是，工作状态3的群体在度假或其他类贷款上有较高的数量，这可能与他们的经济状况和需求有关。

通过以上分析，我们可以得出结论，购车贷款是跨越所有工作状态的普遍需求，而不同工作状态的个体也可能因其特定职业的性质而对某些类型的贷款有更多的需求。



2023-12-21 22:00



Show work

根据执行结果，我们可以得出以下分析结论：

贷款金额统计分析结果：

- 拥有房屋('own')的申请人数最多，共713人，他们的贷款金额的平均值是3060.94（单位：德国马克），标准差是2659.58，中位数是2238.0。
- 租房('rent')的申请人数最少，共179人，但他们的贷款金额的平均值和拥有房屋的申请人差不多，平均值是3122.55，标准差是2553.52，中位数是2301.0。
- 无偿使用房屋('free')的申请人数是108人，平均贷款金额最高，为4906.21，标准差也最大，是3667.00，中位数是3800.5。

信用风险分布分析结果：

- 拥有房屋的申请人中，有527人属于信用风险评级为好('good')，186人为坏('bad')。
- 租房的申请人中，有109人信用风险评级为好，70人为坏。
- 无偿使用房屋的申请人中，64人信用风险评级为好，44人为坏。

关联性分析：

从这些数据可以观察到几个趋势：

- 拥有自己房子的申请人中，信用风险评级为好的比例相对较高，这可能表明房产拥有者通常具有更稳健的财务状况。
- 无偿使用房屋者倾向于申请更高金额的个人贷款，这或许因为他们没有房屋支出，可以负担更高的还款额。
- 租房的申请人的贷款金额平均值和中位数与拥有房屋的申请人相近，但是信用风险评为坏的比例稍高，这可能暗示了租房者相对较高的财务压力或不稳定性。

从这个简单的分析中，我们可以看到房屋拥有情况似乎与贷款金额和信用风险等级有一定的关联。拥有房屋的申请人倾向于获得更好的信用评级，而无偿使用房屋的申请人则申请了更高额的贷款。

以上结论是基于现有数据的统计分析，但需注意这是一种简化的分析和非全面的数据集。要得出更准确的结论，可能需要考虑更多的变量和对数据集进行更深入的分析，例如使用逻辑回归模型来评估不同变量对信用风险评级的影响力。



2023-12-21 22:00



2023-12-21 22:03



Show work

卡方检验的结果显示卡方值 (Chi-squared) 为21.26，自由度 (Degrees of freedom) 为7，p值 (p-value) 为0.0034。由于p值小于0.05，我们可以得出结论，性别分布与贷款目的之间存在显著的关联。

综合上述分析，我们可以回答最初的问题：根据贷款的目的，客户的年龄和性别特征确实有显著差异。年龄在不同贷款目的间存在显著差异，而性别分布也与贷款目的之间存在显著的关联。

这些信息对银行和金融机构来说可能特别有用，因为它们可以帮助这些机构更好地理解他们的客户群体、调整贷款产品以适应特定的客户群体，或者为不同性别和年龄的人群制定针对性的市场营销策略。



2023-12-22 08:05



Show work

根据您提供的代码执行结果，住房所有权与贷款金额之间的相关系数约为-0.1175。这表明两者之间存在轻微的负相关关系，即拥有自己的住房的人通常申请的贷款金额稍微低一些，但这种相关性很弱，不能作为一个强有力的预测或决策依据。

相关系数的绝对值很接近0，意味着住房所有权对贷款金额的影响可能受到其他因素的影响较大，不能简单地断定住房所有权是决定贷款金额高低的主要因素。在实际的信贷决策过程中，还需要结合其他的金融和个人信息来做更全面的分析。

复制

2023-12-22 08:07

储蓄账户、支票账户与信用风险等级的关系如何



2023-12-22 08:09



Show work

分析结果显示：

1. 对于储蓄账户（Saving_accounts）与信用风险等级（Risk）的关联性：

- “little”（少量储蓄）的账户在“good”（信用风险等级好）的情况下数量为386，而在“bad”（信用风险等级坏）的情况下数量为217。
- “moderate”（中等储蓄）的账户在“good”等级下数量为69，在“bad”等级下数量为34。
- 没有储蓄账户记录（none）的在“good”等级下数量为151，在“bad”等级下数量为32。
- “quite rich”（相当富有）的账户在“good”等级下数量为52，在“bad”等级下数量为11。
- “rich”（富有）的账户在“good”等级下数量为42，在“bad”等级下数量为6。

2. 对于支票账户（Checking_account）与信用风险等级的关联性：

- “little”账户在“good”等级下数量为139，在“bad”等级下为135。
- “moderate”账户在“good”等级下数量为164，在“bad”等级下为105。
- 没有支票账户记录（none）的在“good”等级下数量为348，在“bad”等级下为46。
- “rich”账户在“good”等级下数量为49，在“bad”等级下为14。

综合以上分析，可以得出以下结论：

- 在“good”信用风险等级中，拥有较少储蓄（little）的数量最多，其次是中等（moderate）和无储蓄账户记录（none）。富有（rich）和相当富有（quite rich）的情况较少，但在“good”等级中的比例相对于“bad”等级中较高，这意味着储蓄账户中存款多可能与较低的信用风险相关联。

- 在支票账户方面，“none”记录的数量在“good”等级下最多，提示没有支票账户记录的个体可能拥有更好的信用风险等级。而且富有（rich）账户在“good”等级下的数量，虽然绝对值不大，但比例上也超过了“bad”等级。

这些统计数据表明，储蓄账户和支票账户的状态可能和信用风险等级具有一定的相关性。拥有更多储蓄和支票账户的个体倾向于拥有更好的信用等级。当然，为了得出更确定的结论，还需要进一步的统计检验和分析，例如考虑其他可能影响信用风险的变量，以及进行多变量分析。

复制



2023-12-22 08:30



Show work

上述对话涉及了关于个人贷款信息的数据库查询分析。用户使用SQL查询从`personal_loans`表中选择了工作类别（Job）和信用风险等级（Risk）以及各类别计数，接着助理编写了一个Python脚本来连接数据库并执行该查询，最终输出了带有计数的结果。

根据执行的SQL查询结果，得到了不同工作类别和信用风险等级的人数统计，具体情况如下：

- 对于未受技能工作和非居民（Job为0）的人群，有7人的信用风险为`bad`，15人为`good`。
- 对于未受技能工作和常住居民（Job为1）的人群，有56人的信用风险为`bad`，144人为`good`。
- 对于技能工作（Job为2）的人群，有186人的信用风险为`bad`，444人为`good`。
- 对于高技能工作（Job为3）的人群，有51人的信用风险为`bad`，97人为`good`。

通过这个分析，我们可以看出拥有技能的工作者（Job为2）在`good`信用风险评级的数量最多，暗示他们的信用风险较低。同样，未受技能工作的人员（Job为0和1）中`good`评级的比例也比较高。

这些数据对于金融机构有着实际的应用价值，可以帮助它们理解不同工作类别与信用风险之间的关系，从而制定出更加合适的贷款审核策略。

对话结尾，用户和助理都发出了结束对话的信号“”。



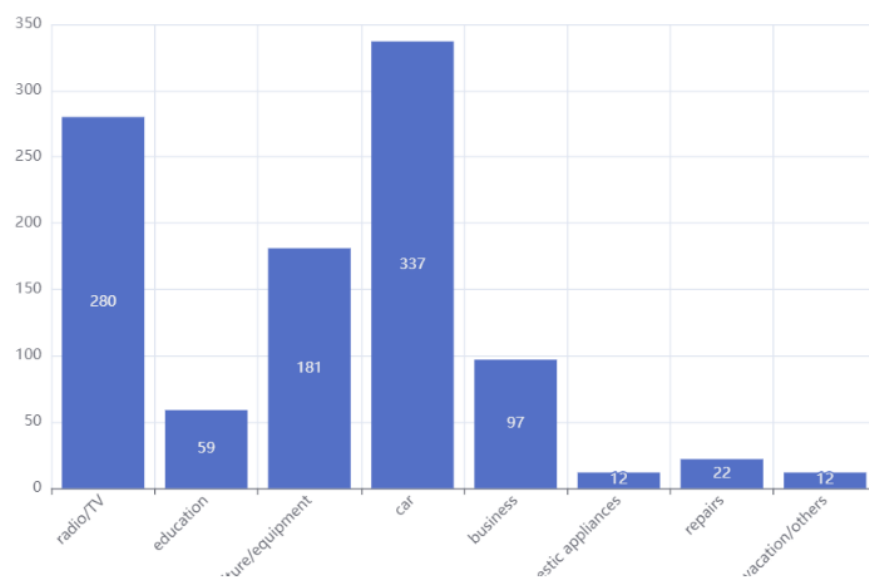
2023-12-22 08:12



Show work

Distribution of Loan Purposes

Loan Purpose Distribution



已成功生成的图表显示了个人贷款表中不同贷款目的分布情况。根据图表，我们可以看到贷款目的主要包括以下几种：电视/广播、教育、家具/设备、汽车、商业、家用电器、维修和假期/其他。各个目的的贷款数量分别为：电视/广播有280笔、教育有59笔、家具/设备有181笔、汽车有337笔、商业有97笔、家用电器有12笔、维修有22笔和假期/其他有12笔。这些数据以条形图的形式呈现，条形图的水平轴代表不同的贷款目的，垂直轴代表相应目的的贷款数量。图表清晰地展现了各个贷款目的的受欢迎程度。

信用风险等级 (Risk) 是否与贷款目的 (Purpose) 之间存在某种相关性



2023-12-22 08:41



Show work

执行结果显示, 卡方检验得出的统计量大约为13.64, 对应的P值大约为0.058。由于这个P值大于0.05, 我们可以得出结论, 在我们设定的显著性水平下, 贷款目的 (Purpose) 和信用风险等级 (Risk) 之间没有统计学意义上的显著相关性。这意味着从我们拥有的数据来看, 不能说贷款目的会显著影响信用风险等级。

请注意, 这并不意味着这两个变量实际上完全独立, 而只是说在现有的数据和分析方法下, 我们无法证明它们之间存在显著的相关性。

2023-12-22 08:13

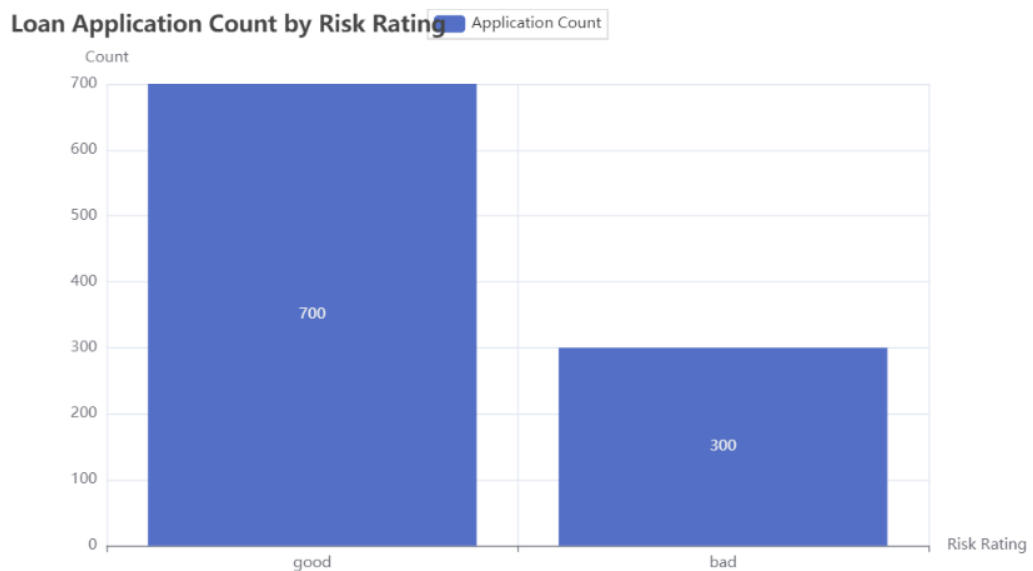
计算不同风险评级下的贷款申请数量分布情况



2023-12-22 08:15



Show work



已生成的图表中显示了两个主要的风险等级——'good'和'bad'对应的贷款申请数量。具体而言, 'good'风险等级的贷款申请数量为700个, 而'bad'风险等级的贷款申请数量为300个。这表明评级为'good'的贷款申请数量多于评级为'bad'的申请, 大约是后者的两倍多。这些数据可以帮助金融机构进行风险控制和贷款产品设计的决策。

复制



2023-12-22 08:18



Show work

从对话中我们可以总结出问题的答案，即关于不同储蓄账户和支票账户类型的贷款申请人数量及其各自的百分比。通过查询`personal_loans`表并计算相关数据，得出以下分析结果：

对于储蓄账户(Saving Accounts)类型：

- 无账户(NA)的贷款申请人有183人，占比18.3%。
- 低储蓄(little)的贷款申请人有603人，占比60.3%。
- 相当富裕(quite rich)的贷款申请人有63人，占比6.3%。
- 富裕(rich)的贷款申请人有48人，占比4.8%。
- 适中(moderate)的贷款申请人有103人，占比10.3%。

对于支票账户(Checking Account)类型：

- 低支票(little)账户的贷款申请人有274人，占比27.4%。
- 适中(moderate)账户的贷款申请人有269人，占比26.9%。
- 无账户(NA)的贷款申请人有394人，占比39.4%。
- 富裕(rich)账户的贷款申请人有63人，占比6.3%。

通过这些数据分析，金融机构可以更加了解客户的储蓄习惯和支票账户情况，从而在审批贷款过程中对风险因素做出更明智的考虑。

复制



2023-12-22 08:22



Show work

根据您的执行结果，已经成功生成了一个名为“Loan Applications by Job Category”的柱状图，该图表展示了不同职业类型的贷款申请数的对比情况。图表的具体数据和设置都包含在了返回的JSON格式的数据中，您可以利用这个JSON数据在前端展现图表。

贷款申请数按职业分类的情况如下所示（注意：“0”代表无专业和非居民工人，“1”代表无专业和居民工人，“2”代表有专业技能，“3”代表高技能工人）：

- 类别“2”有最多的申请数量，共计630个。
- 类别“1”有200个申请。
- 类别“3”有148个申请。
- 类别“0”有22个申请。

这些信息表明，具有专业技能的人群提交了最多的贷款申请。

以上信息可供您用于理解不同工作类型人群的贷款需求，从而对贷款产品的开发和营销策略进行优化。



根据您给出的执行结果，我们可以对不同贷款期限的信用风险情况进行以下分析：

- 短期贷款（少于12个月）中，绝大部分贷款持有者的信用风险为“good”。这暗示这个时段的贷款相对而言风险较小。
- 随着贷款期限增加，拥有“bad”信用风险的人数逐渐上升。例如，贷款期限为12个月，信用风险为“bad”的数量是49，而“good”为130；这表明随着期限增加，不良信用的比例有所上升。
- 当贷款期限为18到24个月时，我们可以看到一个明显的风险增加，其中“bad”的数量为42和56，而“good”的数量则为71和128。这个期限可能是信用风险较高的敏感区域。
- 特别需要关注的是36个月和48个月的贷款期限，因为这两个期限的“bad”信用风险较高（分别为37和28），而且在48个月贷款中，“bad”和“good”的比例接近，数额分别为28和20，这可能表明长期贷款风险显著增高。
- 最长期限（60个月和72个月）的数据样本较少，不过在60个月的贷款中，“good”和“bad”几乎持平，信用风险相对较高。

从这个分析中，我们可以得出结论，随着贷款期限的增加，信用风险呈现出上升的趋势。特别是在18到48个月的贷款中，较长时间的贷款需要更严谨的信用评估。贷款机构在审批较长期贷款时应更加慎重，可能需要加入更多的风险管理措施。

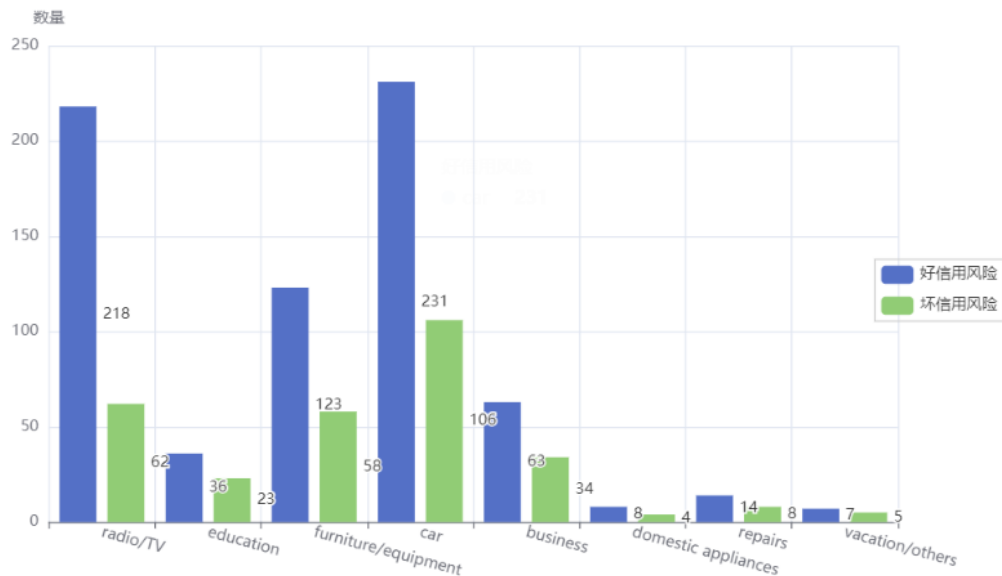


2023-12-22 08:37



Show work

不同贷款目的的信用风险分布



生成的图表名为“不同贷款目的的信用风险分布”，在这个图表中，我们可以看到不同贷款目的与信用风险等级（好或坏）之间的分布情况。图表的X轴分别列出了不同的贷款目的，例如“radio/TV”（电台/电视）、“education”（教育）、“furniture/equipment”（家具/设备）、“car”（汽车）、“business”（商业）、“domestic appliances”（家用电器）、“repairs”（维修）和“vacation/others”（度假/其他）。Y轴显示了数量，表明每个贷款目的下的信用风险等级数量。

在图表中，每个贷款目的都有两种颜色的条形，分别代表好信用风险和坏信用风险。颜色的区分使得人们能直观地识别和比较每一种贷款目的下好信用风险与坏信用风险的数量。通过这个直观的视觉展示，我们可以快速了解到某一特定贷款目的的信用风险状况，以及不同贷款目的之间信用风险的对比。

根据图表的数据，我们可以进行进一步的分析，例如哪些贷款目的伴随的好信用风险较多，哪些则是坏信用风险较多，这样的分析对于理解客户群体的信用行为以及风险管理具有重要作用。

2023-12-22 08:48



复制

2023-12-22 08:48



Show work

根据计算得出的Pearson相关系数，支票账户存款水平 (Checking_account) 与贷款金额 (Credit_amount) 之间的相关系数为0.0246。这个值接近0，表明这两个变量之间的相关性非常弱。也就是说，在我们分析的数据中，支票账户存款水平与贷款金额之间几乎不存在线性关系。

如果有需要进一步分析这两个变量之间的关系，我们可以考虑使用可视化工具（如散点图）或者考虑它们与其他变量的联合关系。但目前得出的结论是没有明显的线性相关性。



2023-12-22 08:39



Show work

根据执行的查询和数据分析，我们可以得知不同职业状态(Job)与储蓄账户情况(Saving_accounts)之间的关系如下：

- 职业状态0（未技术和非居民）：有13个人的储蓄账户状态为'little'，没有人为'moderate'，3人为'quite rich'，1人为'rich'，以及5个未知。
- 职业状态1（未技术和居住者）：有128个人的储蓄账户状态为'little'，23人为'moderate'，12人为'quite rich'，8人为'rich'，以及29个未知。
- 职业状态2（技术工人）：有366个人的储蓄账户状态为'little'，66人为'moderate'，42人为'quite rich'，36人为'rich'，以及120个未知。
- 职业状态3（高技能）：有96个人的储蓄账户状态为'little'，14人为'moderate'，6人为'quite rich'，3人为'rich'，以及29个未知。

分析结果表明，技术工人（Job=2）是储蓄账户情况记录最多的群体，他们的储蓄账户状态以'little'为主，其次是'moderate'和'quite rich'。而高技能工人（Job=3）和未技能的居住者（Job=1）也有较多的记录，但数量少于技术工人，储蓄状态同样以'little'为主，其次顺序为未知、'moderate'、'quite rich'和'rich'。未技术和非居民（Job=0）的记录总体较少。

通过以上数据，企业可以更好地了解不同职业人群的储蓄账户情况，从而在提供贷款服务时作出更合理的决策。

2023-12-22 08:51

分析在不同性别和年龄段的情况下贷款金额的分布情况



2023-12-22 08:52



Show work

根据上述输出，我们可以进行以下分析：

- 在 20-29 岁年龄组中，男性平均贷款金额（3383.27 DM）高于女性（2759.75 DM），男性的总贷款金额和贷款笔数也都高于女性。
- 同样的趋势在 30-39 岁年龄组中也有体现，男性在这个年龄组的平均贷款金额（3479.49 DM）略高于女性，同时总贷款金额和贷款笔数也明显高于女性。
- 40-49 岁年龄组中男性的平均贷款金额和总贷款金额也高于女性，但女性的贷款笔数较少。
- 女性在 50-59 岁和 60-69 岁年龄组的平均贷款金额略低于男性，贷款笔数也明显少于男性。
- 在年龄较大的 70-79 岁组别中，男性平均贷款金额高达 4639.20 DM，远高于女性的 3251.50 DM，但是贷款笔数较少。

需要注意的是，年龄在 10 岁以下以及 80 岁以上的数据都是空值或者为 0，这意味着在此数据集中，我们没有这些年龄段的信息。

总体来说，男性在几乎所有年龄组的平均贷款金额、贷款总额以及贷款笔数都高于女性，特别是在 30-39 岁和 20-29 岁年龄组中贷款活动最为频繁。这可能反映了贷款市场对不同性别和年龄组的贷款需求和服务的不同。



2023-12-22 08:44



Show work

根据您提供的描述性统计信息，我们可以看到性别和住房状况如何影响贷款金额的不同方面。下面是每个组合的贷款金额分析：

1. 女性拥有自己的房屋 (female, own) :

- 平均贷款金额: 2683.56 DM
- 中位数贷款金额: 1812.50 DM
- 总贷款数量: 196
- 最低贷款金额: 250.00 DM
- 最高贷款金额: 18424.00 DM

2. 女性租房 (female, rent) :

- 平均贷款金额: 2820.96 DM
- 中位数贷款金额: 2124.00 DM
- 总贷款数量: 95
- 最低贷款金额: 392.00 DM
- 最高贷款金额: 11590.00 DM

3. 女性无房产 (female, free) :

- 平均贷款金额: 5165.37 DM
- 中位数贷款金额: 3757.00 DM
- 总贷款数量: 19
- 最低贷款金额: 1190.00 DM
- 最高贷款金额: 14782.00 DM

4. 男性拥有自己的房屋 (male, own) :

- 平均贷款金额: 3204.01 DM
- 中位数贷款金额: 2346.00 DM
- 总贷款数量: 517
- 最低贷款金额: 338.00 DM
- 最高贷款金额: 15857.00 DM

5. 男性租房 (male, rent) :

- 平均贷款金额: 3463.64 DM
- 中位数贷款金额: 2414.50 DM
- 总贷款数量: 84
- 最低贷款金额: 276.00 DM
- 最高贷款金额: 15945.00 DM

6. 男性无房产 (male, free) :

- 平均贷款金额: 4850.89 DM
- 中位数贷款金额: 3844.00 DM
- 总贷款数量: 89
- 最低贷款金额: 700.00 DM
- 最高贷款金额: 14318.00 DM

从上面的数据可以看出：

- 通常来说，没有房产的个人（无论男性还是女性）获得的贷款金额平均值和最大值都比拥有房屋的或租房的个人要高。
- 男性和女性之间在相同的住房状况下，男性的平均贷款金额略高于女性的平均贷款金额，尤其在无房产的情况下差异更明显。
- 特别是，女性无房产的贷款金额中位数和最高值均高于其他组合，这可能表明在个别情况下，某些无房产女性可能会有较高的贷款需求。

以上分析提供了一般性的趋势，但请注意这些数值不一定表明性别和住房状况是决定贷款金额的唯一或直接因素。可能还有其他变量影响贷款金额，但没有在当前分析中探讨。

2023-12-22 08:45

不同年龄段和贷款期限之间的关系如何



2023-12-22 08:47

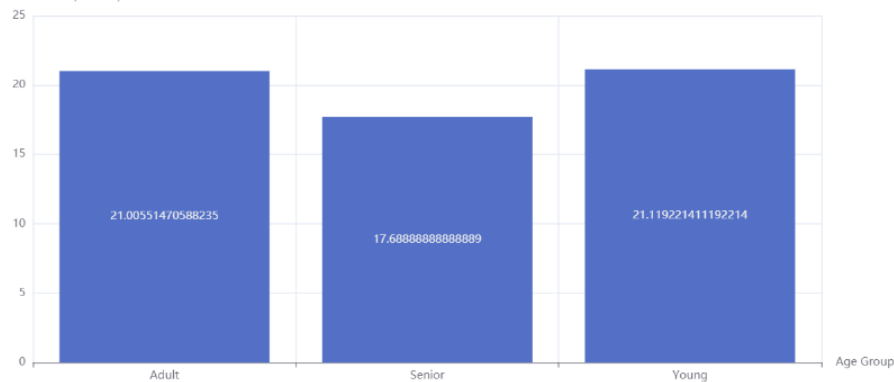


Show work

Average Loan Duration by Age Group

Average Loan Duration

Average Loan Duration (months)



已生成的图表“平均贷款期限与年龄组关系”展示了不同年龄群体的平均贷款期限。根据数据分析结果，我们可以明白以下这些信息：

- 成年人组（31-60岁）的平均贷款期限大约为21.01个月。
- 高龄者组（61岁及以上）的平均贷款期限大约为17.69个月。
- 青年组（18-30岁）的平均贷款期限大约为21.12个月。

从这些数据中可以看出，成年人和青年的平均贷款期限较为相近，而高龄者的贷款期限较短。这可能反映了不同年龄段的消费者在贷款时考虑的因素和偏好有所不同。



2023-12-22 08:59

Show work

根据执行结果，我们现在有了贷款目的为购车的男性和女性申请人的信用风险水平分布数据。结果显示，由数据得出以下分析：

- 男性申请者中，有5名被评为'bad'（不良）风险，10名被评为'good'（良好）风险。
- 女性申请者中，有2名被评为'bad'风险，2名被评为'good'风险。

此分析表明，在此样本中，男性申请者的数量显著大于女性申请者，并且男性申请者中评为良好信用风险的比例也高于不良信用风险。同样地，虽然数据样本较小，女性申请者的良好和不良风险分布是均等的。

请注意，结果的准确性高度依赖于提供的样本数据的代表性和规模。而且，没有进一步的统计分析，我们不应该得出关于总体的任何结论。

如果需要进行更详细的统计分析，我们可以计算风险级别在性别之间分布的概率或百分比，并通过统计检验来确定这些差异是否在统计学意义上是显著的。目前就我们得到的数据来看，这是针对特定样本的初步描述性结果。

以上就是针对这次数据分析的解释和结论。



复制

2023-12-22 10:18

Show work

执行结果显示了为每种性别和账户的组合计算的贷款持续时间的统计信息。具体来说，我们可以看到以下信息：

- 有一位female（女性）申请人拥有little（一点点的）储蓄和moderate（适中的）支票账户，贷款持续时间为48个月。
- 有一位male（男性）申请人没有指定储蓄账户，拥有little（一点点的）支票账户，贷款持续时间为6个月。
- 有两位male（男性）申请人拥有little（一点点的）储蓄和支票账户，贷款持续时间的平均数为33个月，一个最大为42个月，一个最小为24个月。
- 另外，有一位male（男性）申请人拥有little（一点点的）储蓄和moderate（适中的）支票账户，贷款持续时间为12个月。

由于样本数量较小，统计结果可能没有代表整个数据集的趋势。此外，我们发现数据中确实存在缺失值，这些缺失值在分析中被适当处理。

从统计数据中可以看出，不同性别以及不同储蓄和支票账户持有者之间的贷款持续时间分布是有所不同的。女性（只有一个数据点）较高的贷款持续时间可能表明个别情况，而男性在不同储蓄和支票账户组合下显示出贷款持续时间的多样性。

请注意，这仅仅是根据所提供的少量数据样本得出的描述性统计结果。要获得更全面的理解，我们需要一个更大的数据集以及更深入的分析。

至此，我们已经回答了问题，并进行了数据分析。如果您有更具体的要求，例如如果希望看到数据可视化，请告知。否则，任务已完成。



2023-12-22 10:55



Show work

根据卡方检验结果，我们发现在以下贷款用途下，支票账户余额水平对信用风险评级有显著影响：

1. 汽车 (car)
- P-value: 约 $5.80e-11$
2. 家用电器 (domestic appliances)
- P-value: 约0.0384
3. 教育 (education)
- P-value: 约0.00475
4. 家具/设备 (furniture/equipment)
- P-value: 约0.0007318
5. 收音机/电视 (radio/TV)
- P-value: 约 $3.40e-10$

以上P-value都小于0.05，意味着在5%的显著性水平下，我们拒绝了支票账户余额水平与信用风险评级之间无关联的零假设。这表明，对于汽车、家用电器、教育、家具/设备和收音机/电视这些贷款用途，不同的支票账户余额水平对信用风险评级的确有显著影响。

而对于业务 (business)、修理 (repairs)、假期/其他 (vacation/others)，P-value较大，没有统计学意义上的显著差异，表明这些贷款用途下，支票账户余额的不同水平对信用风险评级的影响不显著。

总结来说，支票账户余额水平在某些特定的贷款用途下与信用风险评级有显著的相关性。金融机构在评估贷款申请时，应特别注意这些用途的贷款申请者的支票账户余额水平，因为它可能是判断信用风险的一个重要因素。

2023-12-22 14:16

平均贷款期限是多少？



复制

2023-12-22 14:17



Show work

平均贷款期限已经计算出来了：贷款的平均期限是20.9030个月。

如果您还有其他的问题或数据请求，请告诉我。如果没有的话，任务就此完成。



2023-12-22 10:58

Show work

代码成功执行，我们可以看到在信用风险为"bad"的贷款记录中，按照年龄和贷款期限不同组合得到的分布统计数据。

结果显示了如下的年龄和期限组合的分布：

- 0-25岁组在0-12个月贷款期内的贷款金额分布：有1条记录，平均贷款金额为1295 DM（德国马克），由于只有一笔贷款因此没有标准差。
- 0-25岁组在36-48个月贷款期内的贷款金额分布：有2条记录，平均贷款金额为5129.5 DM，标准差为1161.78 DM。
- 25-35岁组在24-36个月贷款期内的贷款金额分布：有1条记录，贷款金额为5234 DM。
- 45-55岁组在12-24个月贷款期内的贷款金额分布：有1条记录，贷款金额为4870 DM。

这些数据为我们提供了一个初步的视角，通过年龄和贷款期限的组合来看信用风险为"bad"的贷款客户的信用金额分布。由于某些组合中只有一条记录，这些组合的统计数据（如均值、中位数、最小值和最大值）实际上是相同的。

从这个分析过程可以发现，年轻客户在较短贷款期限内借款较少，而0-25岁组在36-48个月内借款金额会有所增加，可能意味着年轻客户在这个年龄段借更高金额的贷款。然而，这仅是基于四条记录的观察，真实的数据集中可能包含更多的记录，因此在分析整体数据时还需要更深入的分析。

注：以上分析基于极其有限的样本数据，只有四条记录，并不代表实际数据库中的整体情况。在实际应用中，应对完整的数据集进行详尽的分析以获得有意义的见解。

复制

2023-12-22 11:18

用饼图展现按月范围分组（如0-12月等）的贷款期限分布

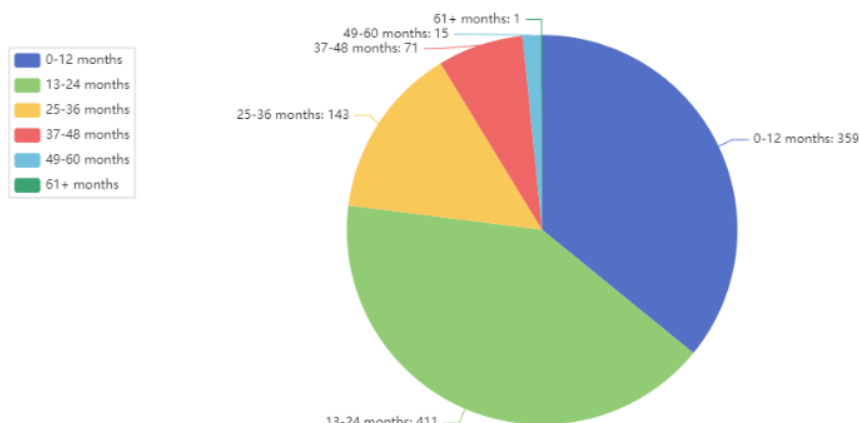


复制

2023-12-22 11:20

Show work

Loan Duration Distribution



已成功生成的饼图展示了不同贷款期限的分布情况。具体数据内容如下：

- 0-12个月：有359笔贷款
- 13-24个月：有411笔贷款
- 25-36个月：有143笔贷款
- 37-48个月：有71笔贷款
- 49-60个月：有15笔贷款
- 61个月以上：有1笔贷款

这些数据反映了贷款持续时间的不同区间段，并通过饼图的形式直观表现出来，从而可以更容易地理解不同期限范围内的贷款分布情况。

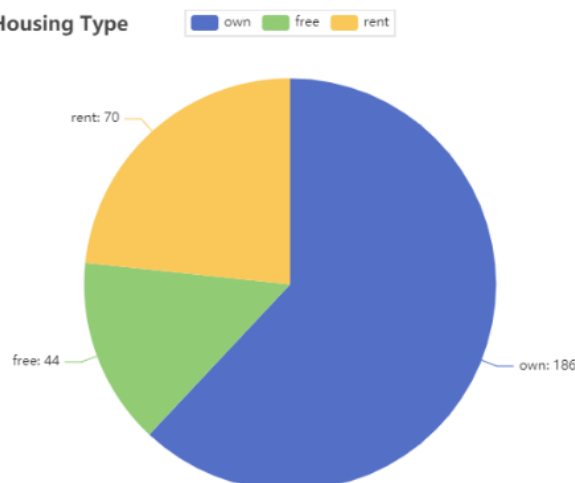


2023-12-22 11:23



Show work

Proportion of Bad Loans by Housing Type



根据生成的图表数据内容，我们可以了解到三种不同住房类型中的不良贷款比例。具体数据如下：

- 自有住房 (own) 的不良贷款占比较高，共有186例。
- 免费住房 (free) 的不良贷款较少，总共有44例。
- 租赁住房 (rent) 的不良贷款也较多，有70例。

通过这份数据，我们可以制作饼图来直观显示自有住房、免费住房和租赁住房在不良贷款中所占的比重，帮助分析并理解不同住房类型与不良贷款之间的关系。

2023-12-22 11:27

用柱状图展现不良贷款-年龄范围&性别占比



2023-12-22 11:28



Show work

Percentage of Bad Loans by Age Range and Gender



已成功生成的图表展示了不同年龄段和性别下，坏账贷款的百分比。以下是图表数据内容的中文介绍：

1. 图表中有两组条形图，分别表示各个年龄段中男性和女性的坏账贷款比例。
2. 年龄段被划分为四个类别：18-30岁、31-45岁、46-60岁，以及61岁以上。
3. 每个年龄段内都有两根柱状图，分别对应男性和女性的数据。
4. Y轴上的百分比数值表示每组中被归类为坏账的贷款比例。

解读结果可知：

- 在18-30岁年龄段中，男性的坏账贷款比例较高；而在31-45岁年龄段中，女性的坏账贷款比例较高。
- 随着年龄的增长，各年龄段的男性和女性的坏账贷款比例都呈下降趋势，其中61岁以上年龄段的坏账贷款比例最低。

这些数据信息对金融机构在贷款审批过程中的风险评估和决策具有重要价值，特别是涉及到年龄和性别人口统计数据。



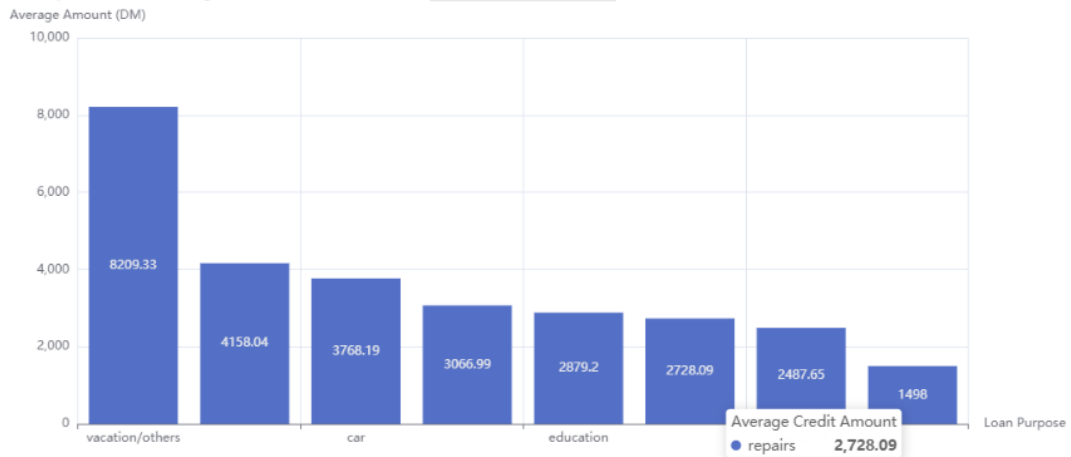
2023-12-22 11:34



Show work

Loan Purpose and Average Credit Amount

Average Credit Amount



根据已生成的图表“Loan Purpose and Average Credit Amount”，我们可以看到不同贷款目的与相应的平均信贷金额的对比。图表中的横轴表示贷款目的，例如“vacation/others”（假期/其他）、“business”（商业）、“car”（购车）等等，而纵轴表示的是平均信贷金额，单位为德国马克（DM）。

从图表中我们可以观察到，不同贷款目的的平均信贷金额存在差异。例如，“vacation/others”目的的平均贷款金额最高，为8209.33 DM；而“domestic appliances”（家电）目的的平均贷款金额最低，仅为1498.00 DM。这样的数据对于银行和贷款机构来说很有价值，可以帮助它们了解不同贷款目的的信贷需求情况，并据此制定相应的信贷政策和产品。

复制



2023-12-22 12:17



Show work

根据分析结果，我们可以看到，贷款目的为“vacation/others”（假期/其他）的借款人在贷款风险中被评为“坏”的比例最高，达到了41.66%。这意味着选择“vacation/others”（假期/其他）目的的借款人更可能被认为是高风险的。

综上所述，目的为“vacation/others”（假期/其他）的贷款最可能导致借款人被判定为高风险。综合数据分析，我们可以向金融机构建议对于此类目的的贷款申请进行更严格的风险评估。



2023-12-22 12:23



Show work

根据您运行的代码，我们得到了一个结果，显示储蓄状态为“little”（少量）的借款人中有50%被认为是“good”风险。然而，这个数据是基于一个非常小的样本来分析的，仅仅4个样本数据点。为了得到一个更有代表性的结论，我们需要分析全部的数据。

由于获取的数据样本很小且不包括所有储蓄状态，我们不能根据这个样本得出结论。理想情况下，我们需要获取更大的数据集并执行相同的分析来观察不同储蓄状态的借款人的信用风险比例。

在现实中，为了回答您的问题（储蓄较多的借款人是否倾向于被视为风险较低），我们需要在包含所有可能的Saving_accounts值的完整数据集上重新运行上面的代码，以得到每个分组的“good”风险比例。如果储蓄较多的组别（如“rich”或“quite rich”）有更高的“good”风险比例，那么可以推断储蓄较多的借款人确实倾向于被认为是风险较低的。

为了进行有效的分析，请在获得完整的数据集之后再次运行分析代码，然后根据得到的结果进行全面的分析。

2023-12-22 12:24

是否存在任何缺失数据或异常值需要进行处理，应该怎样处理



复制

2023-12-22 12:25



Show work

看起来，你提供的信息是关于“personal_loans”数据表的描述，这张表格储存了个人贷款相关的信息。每个字段的“in_use”和“is_pass”都是1，这意味着所有字段都在使用中且通过了某种验证。

请问你具体想对“personal_loans”表执行什么样的任务？是要分析数据、检索信息、还是执行一些特定的查询？给我一些具体的指令，我才能更好地帮助你。

2023-12-22 12:34

分析不同储蓄账户类型客户对于贷款用途的倾向性



2023-12-22 12:36



Show work

根据您执行的代码产生的输出，我们可以看到每个储蓄账户类型与每种贷款目的之间的关联计数。下面是一些基本的分析：

1. “NA”（可能指未指定或没有储蓄账户的情况）类别中，最多的贷款目的是购买汽车（“car”）。
2. 对于“little”储蓄账户的客户来说，最常见的贷款目的也是购买汽车（“car”），其次则是购买家具/设备（“furniture/equipment”）和采购收音机/电视（“radio/TV”）。
3. “moderate”储蓄账户的客户在购买汽车（“car”）方面的贷款需求最高。
4. 对于“quite rich”和“rich”储蓄账户的客户，购买汽车（“car”）和装修（“repairs”）的贷款需求相对较低。

请注意，上述分析只是基于提供数据的计数，不考虑数据集的全面性或客户群的代表性。此外，“NA”类别如何解释取决于数据集是如何处理缺失值的。在进行相关商业决策之前，可能还需要更深入的统计分析和数据验证。

数据已经被保存到“tendency_of_savings_accounts.csv”文件中。这个文件可以用于进一步的分析和数据可视化，为决策提供更丰富的洞察。

如果您需要进一步的分析或有其他问题，请告知。否则，我们的任务已经完成。



2023-12-22 14:19



Show work

根据上述数据，银行更愿意贷款给风险等级为“良好”的客户。我们可以通过对数据的分析，得出银行倾向于贷款给哪类人的一些特点。具体来说：

1. 年龄分布：客户年龄似乎分布广泛。
2. 支票账户余额 (Checking_account)：有不少客户的支票账户信息为'NA'，可能意味着这些客户没有被考虑进账户余额的情况，或者信息不全。但也有客户显示有'little', 'moderate'和'rich'不同程度的余额。
3. 信用额度 (Credit_amount)：信贷金额也是分布广泛，没有特定的偏好。
4. 贷款时长 (Duration)：分布不一，包括从几个月到几十个月的不同贷款期限。
5. 住房状况 (Housing)：大多数的好风险客户都拥有自己的房子 ('own')。
6. 工作状况 (Job)：多样，从非技术、非常住居民 (0) 到高技术熟练工 (3)。
7. 贷款目的 (Purpose)：有多种不同的贷款目的，包括购车、家具/设备、教育等。
8. 风险等级 (Risk)：所有被分析的样本风险等级都是'good'，意味着它们是银行更愿意贷款给他们的样本。
9. 储蓄账户 (Saving_accounts)：同支票账户余额一样，有不少客户的储蓄账户信息为'NA'，说明信息不全。
10. 性别 (Sex)：既有男性也有女性。

综上所述，虽然我们能观察到一些特点，例如大多数好风险的客户都有自己的房产，但由于数据集中具体数值的多样性，很难总结出绝对的规则。此外，'NA'可能表示数据不完整，这也可能会影响分析结果。

为了更深入地分析这些属性与被认为是低风险贷款（即银行更愿意贷款的对象）之间的关系，可以使用统计检验或机器学习模型来识别哪些属性对于预测客户是否被视为良好风险最为重要。然而，基于当前提供的数据，我们可以保守地说，拥有自己房子的客户、拥有储蓄账户的客户和各种年龄和工作情况的客户都有机会被视为良好风险。因此，我们不能明确指出银行更倾向于贷款给某一具体类别的人。

帮我从不同角度分析，提出10个不同的问题



2023-12-22 17:28



Show work

1. 在 `personal_loans` 表中，平均信用额度 (Credit_amount) 随年龄 (Age) 增长的趋势是怎样的？
2. 不同性别 (Sex) 在不同房屋状况 (Housing) 下的信用风险 (Risk) 分布情况如何？
3. 申请个人贷款的目的 (Purpose) 最常见的是什么，它们的平均信用期限 (Duration) 是多少？
4. 工作类型 (Job) 与信用风险 (Risk) 之间是否存在明显的关联？
5. 在拥有不同储蓄账户状态 (Saving_accounts) 的人群中，哪一类的平均信用额度 (Credit_amount) 最高？
6. 不同的支票账户状态 (Checking_account) 是否影响信用风险 (Risk) 的评定？
7. 哪一年龄段的人更倾向于租房 (Housing=rent)，他们的平均信用额度 (Credit_amount) 和期限 (Duration) 是多少？
8. 教育贷款 (Purpose=education) 是否通常有更低的坏账风险 (Risk=bad)？
9. 信用风险较高 (Risk=bad) 的贷款请求中，最常见的贷款目的 (Purpose) 是什么？
10. 平均而言，哪种类型的职业 (Job) 借款人的信用额度 (Credit_amount) 更高？

