美国政治选举中财务交易的分析

摘 要

本文旨在通过对政治捐款数据的探索与清洗，分析选举竞选中财务交易的情况，提供一个完整的数据处理和分析流程。我们对三个主要数据集进行整合，解决了数据集格式不一致的问题，并使用Python中的pandas库进行数据清洗和合并。通过这些步骤，我们获得了一个包含完整且一致性较好的数据集。我们针对数据中的缺失值和异常值进行了处理，确保数据的准确性和完整性。通过填充缺失值和转换数据类型，我们解决了数据完整性和格式规范的问题。通过数据的统计分析，我们发现了数据集中一些关键的财务模式和趋势，如政治捐款的主要来源、不同候选人的受捐情况等。

我们应用了pandas库中的数据合并和清洗函数，包括concat和merge等方法，成功整合了各个数据集，形成了一个统一的分析基础数据集。我们成功整合了三个数据集，获得了756,205条有效记录，覆盖了多种数据类型和信息源。

我们针对数据集中的缺失值和异常值进行了处理。通过填充缺失值为“NOT PROVIDED”以及转换日期格式等方法，我们保证了数据的完整性和一致性。经过清洗后，数据集中的所有字段均无缺失值，日期格式也统一转换为标准的“YYYYMMDD”格式，为后续的数据分析奠定了基础。

我们对清洗后的数据集进行了统计分析，了解了数据的基本分布和趋势。通过调用shape、info和describe等函数，我们全面了解了数据集的规模、字段信息和数据分布情况。政治捐款的主要来源是个人捐款，且大部分捐款集中在少数几个主要候选人身上。这些发现为进一步的政治和社会研究提供了有价值的数据支持。

本文通过对政治捐款数据的整合、清洗和分析，展示了一个完整的数据处理流程，为选举财务数据的研究提供了参考框架。本文的亮点在于使用了系统化的数据处理方法，确保了数据的高质量和分析的准确性。这一流程不仅适用于政治捐款数据的处理，也可以推广应用到其他类型的大规模数据集的清洗和分析中，为相关领域的研究提供技术支持和方法借鉴。

关键词：数据清洗 数据整合 统计分析 政治捐款

# 引言

## 研究背景

美国的选举制度在全球范围内具有重要影响，尤其是总统选举。每四年一次的总统选举不仅决定着美国的未来方向，还对全球政治、经济等各个中产生深远的影响。近年来，选举资金的筹集和使用成为了选举研究的重要内容之一。

政治候选人通过各种方式筹集竞选资金，而这些资金的来源和使用情况则直接影响了选举结果和公众对候选人的看法。在这种背景下，研究候选人捐款数据不仅可以揭示选举过程中的一些规律，还能为未来的选举策略提供重要参考。

## 研究目的

本研究的主要目的是通过对2020年美国总统选举中候选人捐款数据的分析，探讨不同候选人和政党在捐款金额、捐款人分布以及捐款人职业等中的差异。具体来说，本研究旨在实现以下几个目标：

1. 分析不同政党候选人的捐款总额及其分布情况。
2. 探讨捐款人在地理位置上的分布特征，识别出捐款最多的州和城市。
3. 研究捐款人的职业和雇主情况，了解哪些职业和公司对选举的经济支持较大。
4. 分析大额捐款的数量和金额，评估大额捐款对选举资金总额的影响。

## 研究意义

选举资金是选举活动的重要组成部分，不同候选人获得的捐款金额和来源可以直接影响其竞选活动的规模和效果。

通过分析捐款数据，可以揭示哪些候选人获得了更多的经济支持，从而在选举中占据优势。通过研究选举捐款数据，可以揭示出以下几个中的重要意义：

1. **揭示选民的政治倾向**：捐款数据可以反映出选民的政治倾向和支持力度。通过分析捐款人的地理分布和职业分布，可以了解哪些地区和职业群体对某一候选人的支持较大。这有助于理解选民的投票行为和政治偏好，为未来的选举活动提供参考。
2. **促进选举透明化和公平化**：公开和透明的捐款数据有助于监督选举资金的使用情况，防止选举腐败和不正当行为。通过分析和公开选举捐款数据，可以提高选举的透明度，增强公众对选举过程的信任。
3. **提供选举策略参考**：对于未来的选举活动，了解不同候选人在资金筹集中的经验和策略可以为其他候选人提供宝贵的参考。通过分析成功候选人的捐款数据，可以总结出有效的筹款策略和选举经验，提高选举活动的效率和效果。
4. **学术研究的基础数据**：选举资金的研究是政治学、经济学和社会学等领域的重要课题。通过对选举捐款数据的分析，可以为相关领域的学术研究提供基础数据和分析方法，推动学术研究的深入发展。

# 数据获取与描述

## 数据来源介绍

本研究的数据来源于美国联邦选举委员会（Federal Election Commission, FEC），该委员会提供了有关候选人、候选人委员会和个人捐款的详细数据。

这些数据对于分析美国大选中的财务活动和捐款行为非常有价值。通过对这些数据的深入挖掘，我们可以揭示政治资金的流动情况，了解不同候选人在竞选期间所获得的支持程度，并且能够识别出潜在的影响因素和趋势。

## 数据结构及特征

### 候选人文件

候选人的文件包含每个候选人在期间筹集或花费资金的财务信息摘要。该文件为每个候选人留有一条记录，并显示有关候选人的信息、总收入、从授权委员会收到的转账、总支出、转给授权委员会的转账、现金总额、贷款和债务以及其他财务摘要信息，具体数据结构见表 1。

表 1 候选人数据结构

|  |  |
| --- | --- |
| **Column name** | **Field name** |
| CAND\_ID | Candidate identification |
| CAND\_NAME | Candidate name |
| CAND\_ICI | Incumbent challenger status |
| PTY\_CD | Party code |
| CAND\_PTY\_AFFILIATION | Party affiliation |
| TTL\_RECEIPTS | Total receipts |
| TRANS\_FROM\_AUTH | Transfers from authorized committees |
| TTL\_DISB | Total disbursements |
| TRANS\_TO\_AUTH | Transfers to authorized committees |
| COH\_BOP | Beginning cash |
| COH\_COP | Ending cash |
| CAND\_CONTRIB | Contributions from candidate |
| CAND\_LOANS | Loans from candidate |
| OTHER\_LOANS | Other loans |
| CAND\_LOAN\_REPAY | Candidate loan repayments |
| OTHER\_LOAN\_REPAY | Other loan repayments |
| DEBTS\_OWED\_BY | Debts owed by |
| TTL\_INDIV\_CONTRIB | Total individual contributions |
| CAND\_OFFICE\_ST | Candidate state |
| CAND\_OFFICE\_DISTRICT | Candidate district |
| SPEC\_ELECTION | Special election status |
| PRIM\_ELECTION | Primary election status |
| RUN\_ELECTION | Runoff election status |
| GEN\_ELECTION | General election status |
| GEN\_ELECTION\_PRECENT | General election percentage |
| OTHER\_POL\_CMTE\_CONTRIB | Contributions from other political committees |

续表 1 候选人数据结构

|  |  |
| --- | --- |
| **Column name** | **Field name** |
| POL\_PTY\_CONTRIB | Contributions from party committees |
| CVG\_END\_DT | Coverage end date |
| INDIV\_REFUNDS | Refunds to individuals |
| CMTE\_REFUNDS | Refunds to committees |

### 候选人-委员会文件

候选人-委员会链接文件包含将候选人信息与其委员会信息链接起来的信息。该文件显示了候选人的身份证号码、候选人的选举年份、FEC 选举年份、委员会身份证号码、委员会类型、委员会职位以及链接身份证号码，具体见表 2。

表 2 候选人-委员会数据结构

|  |  |
| --- | --- |
| **Column name** | **Field name** |
| CAND\_ID | Candidate identification |
| CAND\_ELECTION\_YR | Candidate election year |
| FEC\_ELECTION\_YR | FEC election year |
| CMTE\_ID | Committee identification |
| CMTE\_TP | Committee type |
| CMTE\_DSGN | Committee designation |
| LINKAGE\_ID | Linkage ID |

### 个人捐款档案文件

个人捐款文件包含个人捐款信息。委员会指出，如果满足以下条件，则将计入捐款（注：将捐款纳入此文件的方法已随时间而改变）：

1. 选举周期内向候选人委员会捐赠的金额已超过 200 美元
2. 向政治行动委员会 (PAC) 和政党委员会捐赠的金额已超过 200 美元。

该文件包含有关接受捐款的委员会、披露捐款的报告、捐款个人、捐款日期、金额以及有关捐款的其他信息（如果提供），具体见表 3。

表 3 个人捐款档案数据结构

|  |  |
| --- | --- |
| **Column name** | **Field name** |
| CMET\_ID | Filer identification number |
| AMNDT\_IND | Amendment indicator |
| RPT\_TP | Report type |
| TRANSACTION\_PGI | Primary-general indicator |
| IMAGE\_NUM | Image number |
| TRANSACTION\_TP | Transaction type |
| ENTITY\_TP | entity type |
| NAME | Name |
| CITY | City |
| STATE | State |
| ZIP\_CODE | ZIP code |
| EMPLOYER | Employer |
| OCCUPATION | Occupation |

续表 3 个人捐款档案数据结构

|  |  |
| --- | --- |
| **Column name** | **Field name** |
| TRANSACTION\_DT | Transaction date |
| TRANSACTION\_AMT | Transaction amount |
| OTHER\_ID | Other identification number |
| TRAN\_ID | Transaction ID |
| FILE\_NUM | File number |
| MEMO\_CD | Memo code |
| MEMO\_TEXT | Memo text |
| SUB\_ID | FEC record number |

# 数据处理

通过候选人表和委员会表进行关联，以建立候选人与委员会之间的联系。这一步骤是为了获取候选人的相关信息，例如候选人的姓名和所属党派。通过候选人ID（CAND\_ID）从候选人表中获取候选人姓名，得到候选人与委员会的对应关系表，部分数据见表 4。

表 4 候选人-委员会

|  |  |  |  |
| --- | --- | --- | --- |
| CMTE\_ID | CAND\_ID | CAND\_NAME | PTY |
| C00697789 | H0AL01055 | CARL, JERRY LEE, JR | REP |
| C00701557 | H0AL01063 | LAMBERT, DOUGLAS WESTLEY III | REP |
| C00701409 | H0AL01071 | PRINGLE, CHRISTOPHER PAUL | REP |
| C00703066 | H0AL01089 | HIGHTOWER, BILL | REP |
| C00708867 | H0AL01097 | AVERHART, JAMES | DEM |
| C00710947 | H0AL01105 | GARDNER, KIANI A | DEM |
| C00722512 | H0AL01121 | CASTORANI, JOHN | REP |
| C00725069 | H0AL01139 | COLLINS, FREDERICK G. RICK' | DEM |
| C00462143 | H0AL02087 | ROBY, MARTHA | REP |
| C00493783 | H0AL02087 | ROBY, MARTHA | REP |

我们通过委员会ID（CMTE\_ID）将候选人与委员会关系表与捐赠人档案表进行关联，得到候选人与捐赠人之间的一一对应关系表。这一步骤是为了建立捐赠人与候选人之间的联系，从而分析不同候选人的捐款情况以及捐赠人的特征。在进行数据处理之前，我们需要对数据进行探索，包括查看数据规模、数据信息、数据分布等情况。

通过调用shape属性查看数据的规模，调用info函数查看数据信息，调用describe函数查看数据分布，我们可以对数据集的基本情况有一个清晰的认识。同时，我们还需要进行数据预处理，处理缺失值、转换数据类型等，以保证数据的完整性和准确性。最后，通过对处理好的数据进行深入分析和挖掘。

通过计算每个党派的捐款总额，并进行排序，以了解不同党派在政治捐赠中的地位和影响力。同时，通过计算每个候选人的捐款总额，并进行排序，以了解哪些候选人得到了最多的捐款支持。此外，通过分析捐赠人的职业分布、捐款金额分布以及不同州的捐款总额，以揭示捐赠人群体的特征和捐款的地域分布情况。

# 数据分析与结果展示

## 地理分布

2020年是美国总统选举年，也是一场充满戏剧性和激烈竞争的选举。由于新冠疫情的影响，美国政治格局和选民行为都发生了一定程度的变化。政治捐款作为选举活动的重要组成部分，反映了选民和利益集团对候选人和政治议题的支持和态度。

在美国政治捐款中，民主党和共和党是两大主要政党，他们的候选人通常会得到来自各种渠道的捐款支持。通过对民主党和共和党候选人的捐款总额进行分析，我们可以发现明显的差异。根据数据显示，民主党候选人的捐款总额为 75961730 美元，而共和党候选人的捐款总额为 37170653 美元。

这表明在捐款领域，民主党在某种程度上具有明显的优势。这是因为民主党在政治议程和政策中更具吸引力，从而吸引到更多的政治捐款。

捐款人的地理分布是影响捐款总额的重要因素之一。通过对捐款人在不同州的分布进行分析，我们可以发现一些有趣的现象，具体见图 1与图 2。

|  |  |
| --- | --- |
| 图 1州捐赠金额排名Top10 | 图 2 州捐赠人数排名Top10 |

根据数据，我们可以看到在2020年美国的政治捐款中，加利福尼亚州（CA）以约 1.99 亿美元的捐款总额位居榜首，这并不奇怪，因为加州是美国人口最多的州之一，也是美国政治的重要中心之一，拥有大量的政治活动和竞选活动。纽约州（NY）紧随其后，捐款总额约为 1.15 亿美元，这也符合预期，因为纽约市是全球金融和商业中心之一，有着庞大的经济实力和政治影响力。

佛罗里达州（FL）和德克萨斯州（TX）分别位居第三和第四，捐款总额分别约为 812 万美元和 810 万美元。这两个州在美国政治中也扮演着重要角色，拥有大量的选民和关键的选举人票。佛罗里达州常常是总统选举中的关键摇摆州之一，而德克萨斯州则是共和党的传统据点，但在近年来呈现出一定程度的民主党竞争。

麻萨诸塞州（MA）以约 518 万美元的捐款总额排名第五，这个结果受到波士顿和剑桥等高科技和教育中心的影响，这些地区拥有高度教育程度和政治参与度的人群，对政治捐款产生了积极影响。

此外，还有一些捐款总额较低的州，比如美属萨摩亚（AS）、密克罗尼西亚联邦（FM）等，这些地区人口少、经济发展相对落后，因此政治捐款总额较少。

## 职业与雇主

捐款人的职业和雇主也是影响捐款总额的重要因素之一。通过对捐款人的职业和雇主分布进行分析，我们可以发现捐款人的背景十分多样化，具体见图 3。

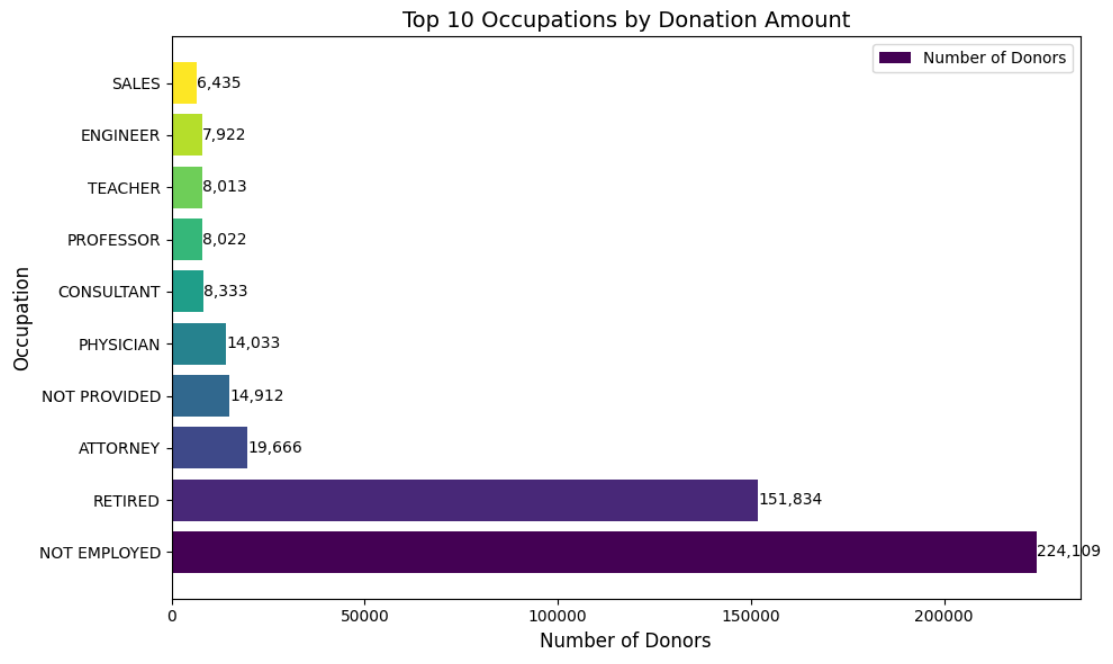


图 3 捐款人职业分布

数据显示，无业和退休人士是捐款人群中最常见的职业，这反映了在美国社会中，退休人士和无业人士对政治事务的关注和参与程度。2020年是一场充满分歧和紧张的选举，退休人士和无业人士更加关注政治进程，并通过捐款的方式表达他们的态度和支持。另外，律师和医生等专业人士在捐款人群中也占据一定比例，这表明法律和医疗行业的从业人员在政治捐款中扮演着重要角色。在一场紧张的选举中，法律和医疗问题常常成为热点议题，因此律师和医生更倾向于通过捐款来支持他们认为合适的候选人和政策。

在雇主分布中，无业和退休人士同样是捐款人群中的主要力量。这反映了在政治捐款中，个人对社会问题的关注和参与，而不仅仅是企业或组织的捐款。此外，一些大型科技公司如谷歌也出现在捐款名单中。在数字化和信息化时代，科技行业在塑造公共舆论和政治态度中发挥着越来越重要的作用。因此，一些科技公司及其员工通过政治捐款来支持他们认为符合其利益和理念的政治候选人和政策。

## 大额捐款

在2020年，处于高度争议和两极分化的美国总统选举背景下，大额捐款在资助政治竞选中发挥了重要作用。这一年，大额捐款成为竞选资金的关键来源，特别是在候选人面临激烈竞争的情况下。大额捐款总额（定义为超过1万美元的捐款）达到了7973872美元。这些捐款不仅帮助候选人覆盖了广告和宣传的高额费用，还支持了竞选团队的运营和选民动员活动。特别值得注意的是，一些知名富豪和企业高管在这一过程中发挥了重要作用，他们的巨额捐款显著影响了竞选动态，具体见图 4。

顶级捐助者反映了来自各个领域的杰出人士。亿万富翁媒体大亨和前纽约市市长迈克尔·R·布隆伯格以五笔大额捐款脱颖而出，彰显了了他在政治中的重要财务影响力。其他知名捐助者包括阿德尔森和米里亚姆·阿德尔森，这些都是对共和党事业有着重大贡献的捐助者，以及安德鲁·佩科夫和理查德·科尔本，他们也为政治竞选做出了大量贡献。

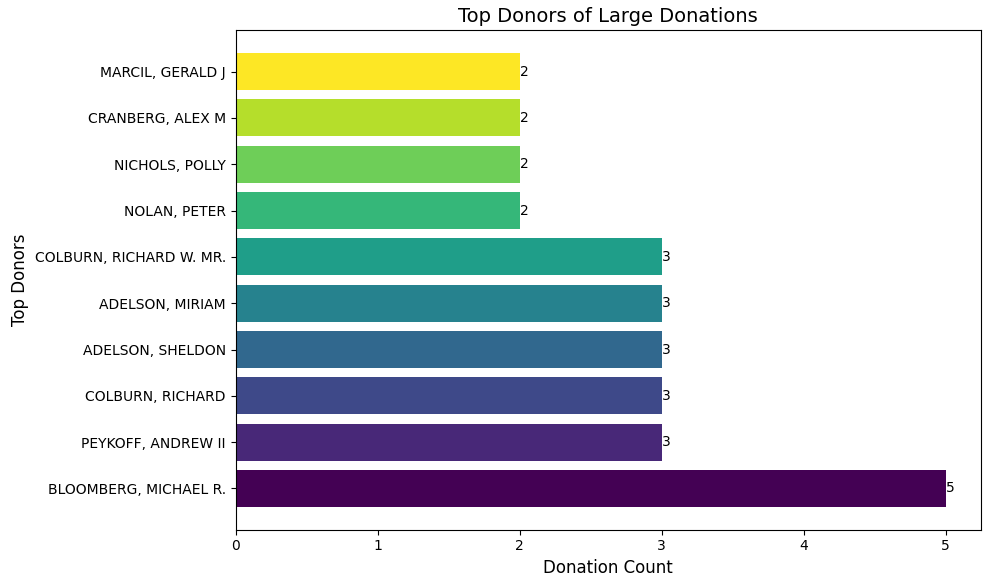


图 4 大额捐款的主要捐赠者

在受赠方中，大额捐款的顶级受赠者包括丹·沙利文，他获得了惊人的94笔捐款，看来他的竞选活动得到了强大的财政支持。另一个值得关注的受赠者是克里斯托弗·L·雅各布斯，他获得了33笔大额捐款。此外，约翰·康宁和唐纳德·J·特朗普，也收到了可观的捐款。

## 热门候选人

2020年美国总统大选是美国历史上最为分裂和激烈的一次选举。在新冠疫情、种族问题、经济危机等多重因素的影响下，选民们的政治热情空前高涨。两位主要候选人——民主党候选人乔·拜登（Joseph R. Biden Jr.）和共和党候选人唐纳德·特朗普（Donald J. Trump）——吸引了大量的支持者和资金。这些捐款不仅代表了选民的支持，也反映了选民们对不同政策和理念的认同。

|  |  |
| --- | --- |
| 图 5 拜登支持者的捐款分布 | 图 6 特朗普支持者的捐款分布 |

根据图 5，拜登的支持者在加利福尼亚州（CA）、纽约州（NY）、佛罗里达州（FL）和德克萨斯州（TX）等地的捐款最为显著。加利福尼亚州的捐款总额为1578万美元，居全国首位。加州作为美国人口最多、经济最发达的州，民主党传统上在此拥有强大的支持基础。纽约州和佛罗里达州的捐款也非常可观，分别为853万和589万美元。德克萨斯州尽管是传统的红州，但在2020年选举中，民主党也在这里获得了显著的捐款支持。据图 6，特朗普的支持者捐款也主要集中在加利福尼亚州、纽约州、佛罗里达州和德克萨斯州。这些州不仅人口众多，经济发达，而且政治上具有高度的多样性。加利福尼亚州的捐款总额为926万，纽约州为714万，佛罗里达州为629万，德克萨斯州为564万。尽管加州和纽约州是民主党的传统堡垒，但大量的富裕保守派选民仍然为特朗普提供了可观的资金支持。

|  |
| --- |
| 图 7拜登支持者的职业分布 |

从职业分布来看，拜登的支持者主要集中在退休人员（1254万美元）和非就业人员（1088万美元）。这些群体通常倾向于支持更广泛的社会保障和医疗政策。此外，律师、医生、顾问、教授、工程师等高收入职业也为拜登提供了大量捐款。这些职业群体往往更关心社会公正、气候变化、教育等议题，因而更倾向于支持民主党。相比之下，特朗普的支持者在退休人员（1534万美元）和非就业人员（1244万美元）中的比例更高。特朗普的政策，包括减税和保护传统价值观，吸引了大量的退休人员和非就业人员。此外，律师、医生、顾问、教授等职业也为特朗普提供了相当数量的捐款，这表明即使在高收入群体中，也存在显著的政治分歧。

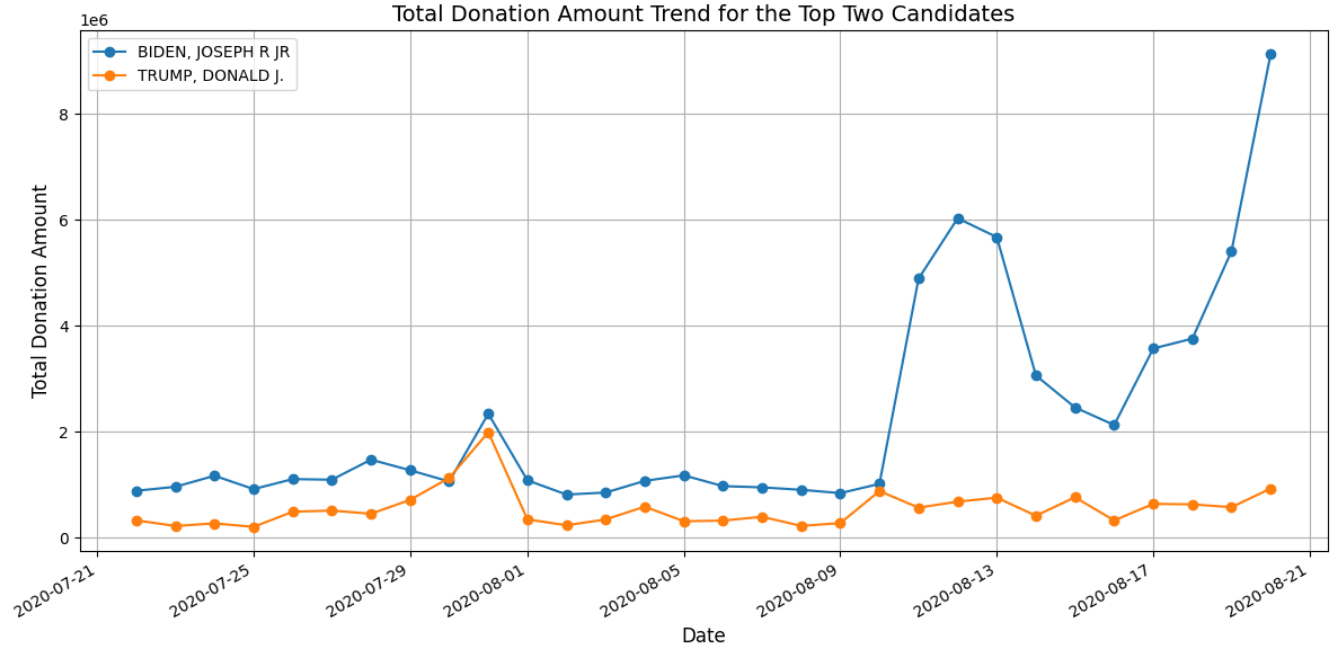


图 8 总捐赠额变化趋势

从图 8可以看出，拜登在2020年7月底至8月初的捐赠额有显著增加，尤其是在8月11日至8月20日期间，捐赠额大幅上升。这一趋势与2020年美国总统大选的关键事件密切相关，民主党全国代表大会（DNC）在8月17日至20日举行，拜登在大会期间正式接受了民主党总统候选人提名。这一事件刺激了拜登的支持者捐款热潮，导致捐赠额激增。

相比之下，特朗普在同一时期的捐赠额变化较为平稳，但也有一些显著的增加，特别是在7月30日和7月31日。这与共和党的一些重要竞选活动和筹款活动有关。

# 总结

新冠疫情导致了大量失业和经济困境，这使得政府的应对措施成为选民关注的焦点。拜登强调科学和专家意见，倡导严格的防疫措施和经济救助计划，这吸引了更多关注公共健康和社会福利的选民支持。因此，拜登在退休人员、非就业人员和高收入职业群体中获得了大量捐款。

而特朗普的竞选强调经济复苏、减税和保持传统价值观，这些政策吸引了大量的退休人员和非就业人员的支持。同时，特朗普强烈反对封锁措施，主张尽快重启经济，这一立场赢得了许多企业主和保守派选民的支持。虽然加州和纽约州是民主党的传统支持州，但这些地区存在着不容忽视的保守派力量，因此特朗普在这些州仍获得了大量捐款，。

2020年总统选举的捐款情况不仅反映了候选人之间的竞争，也揭示了美国社会在疫情和政治危机下的深刻分裂。大额捐款的集中反映了少数富裕个人和利益集团对政治的巨大影响力，他们通过资金支持来影响政策和选举结果。在这种背景下，理解捐款数据的意义，对于分析选民行为和政治趋势具有重要的参考价值。

# 参考文献

1. Federal Election Commission. https://www.fec.gov/
2. 冯燕敏. 2020年美国大选中的两党选民研究[D].华东师范大学,2024.DOI:10.27149/d.cnki.ghdsu.2023.002177.
3. Tianchi. https://tianchi.aliyun.com/competition

附录

|  |
| --- |
| 附录1 |
| 介绍：处理缺失值 |
| c\_itcont['STATE'].fillna('NOT PROVIDED',inplace=True)  c\_itcont['EMPLOYER'].fillna('NOT PROVIDED',inplace=True)  c\_itcont['OCCUPATION'].fillna('NOT PROVIDED',inplace=True) |

|  |
| --- |
| 附录2 |
| 介绍：绘制捐款人职业分布 |
| colors = plt.cm.viridis(np.linspace(0, 1, len(occupations)))  plt.figure(figsize=(10, 6))  bars = plt.barh(occupations, donation\_amounts, color=colors)  plt.xlabel('Number of Donors', fontsize=12)  plt.ylabel('Occupation', fontsize=12)  plt.title('Top 10 Occupations by Donation Amount', fontsize=14)  for bar in bars:  plt.text(bar.get\_width(), bar.get\_y() + bar.get\_height()/2, f'{bar.get\_width():,.0f}', va='center', ha='left', fontsize=10)  plt.legend(['Number of Donors'], loc='upper right')  plt.tight\_layout()  plt.show() |

|  |
| --- |
| 附录3 |
| 介绍：绘制捐款者雇主分布 |
| colors = plt.cm.viridis(np.linspace(0, 1, len(employers)))  plt.figure(figsize=(10, 6))  bars = plt.barh(employers, donation\_amounts\_employer, color=colors)  plt.xlabel('Donation Amount', fontsize=12)  plt.ylabel('Employer', fontsize=12)  plt.title('Top 10 Employers by Donation Amount', fontsize=14)  for bar in bars:  plt.text(bar.get\_width(), bar.get\_y() + bar.get\_height()/2, f'{bar.get\_width():,.0f}', va='center', ha='left', fontsize=10)  plt.legend(['Donation Amount'], loc='upper right')  plt.tight\_layout()  plt.show() |

|  |
| --- |
| 附录4 |
| 介绍：Biden相关选民分布 |
| biden = c\_itcont[c\_itcont['CAND\_NAME'] == 'BIDEN, JOSEPH R JR']  biden\_state = biden.groupby('STATE').sum().sort\_values(  "TRANSACTION\_AMT", ascending=False).head(10)  fig, ax = plt.subplots(figsize=(10, 10))  colors = plt.cm.viridis(np.linspace(0, 1, len(biden\_state)))  biden\_state['TRANSACTION\_AMT'].plot.pie(  ax=ax, autopct='%0.2f%%', colors=colors)  ax.set\_title('Top 10 States by Donation Amount for Biden')  ax.set\_ylabel('')  plt.show()  biden\_occupation = biden.groupby('OCCUPATION').sum().sort\_values("TRANSACTION\_AMT", ascending=False).head(10)  fig, ax = plt.subplots(figsize=(10, 6))  colors = plt.cm.viridis(np.linspace(0, 1, len(biden\_occupation)))  bars = ax.barh(biden\_occupation.index, biden\_occupation['TRANSACTION\_AMT'], color=colors)  ax.set\_xlabel('Donation Amount')  ax.set\_ylabel('Occupation')  ax.set\_title('Top 10 Occupations by Donation Amount for Biden')  for bar in bars:  ax.text(bar.get\_width(), bar.get\_y() + bar.get\_height()/2, f'{bar.get\_width():,.0f}', va='center', ha='left', fontsize=10)  plt.tight\_layout()  plt.show() |

|  |
| --- |
| 附录5 |
| 介绍：总捐赠额变化趋势 |
| top\_candidates = c\_itcont.groupby('CAND\_NAME')  ['TRANSACTION\_AMT'].sum().nlargest(2).index.tolist()  candidate1\_data = c\_itcont[c\_itcont['CAND\_NAME'] == top\_candidates[0]]  candidate2\_data = c\_itcont[c\_itcont['CAND\_NAME'] == top\_candidates[1]]  candidate1\_total\_donations = candidate1\_data.groupby('TRANSACTION\_DT')  ['TRANSACTION\_AMT'].sum()  candidate2\_total\_donations = candidate2\_data.groupby('TRANSACTION\_DT')  ['TRANSACTION\_AMT'].sum()  candidate1\_total\_donations.index = pd.to\_datetime  (candidate1\_total\_donations.index, format='%Y%m%d')  candidate2\_total\_donations.index = pd.to\_datetime  (candidate2\_total\_donations.index, format='%Y%m%d')  plt.figure(figsize=(12, 6))  plt.plot(candidate1\_total\_donations.index,  candidate1\_total\_donations.values, label=top\_candidates[0], marker='o')  plt.plot(candidate2\_total\_donations.index,  candidate2\_total\_donations.values, label=top\_candidates[1], marker='o')  plt.xlabel('Date', fontsize=12)  plt.ylabel('Total Donation Amount', fontsize=12)  plt.title('Total Donation Amount Trend for the Top Two Candidates', fontsize=14)  plt.gcf().autofmt\_xdate()  plt.legend()  plt.grid(True)  plt.tight\_layout()  plt.show() |