# FIN 557 Evaluation of Initial Public Offerings (IPOs) and Market Performance

Group 14

Authors: Akshaya Suresh, Chubo Wen, Preethi Peelamedu Surendran, Purva Vaswani, Vishwas Rao, Xingyu Chen, Yu Bai, Ziqian Ding

## 1) Introduction

In the realm of finance, Initial Public Offerings (IPOs) serve as pivotal events marking a company's transition from private to public ownership. Understanding the intricacies of IPOs and their subsequent market performance is paramount for investors, analysts, and policymakers alike. This analysis endeavors to delve deep into the world of IPOs, employing a multifaceted approach to dissect data, uncover trends, and draw actionable insights.

**Question:** Our primary inquiry revolves around the analysis of Initial Public Offerings (IPOs) and their impact on market performance. Specifically, we aim to explore the dynamics surrounding IPOs, understand their implications on various market sectors, and identify trends in stock price evolution post-IPO.

**Approach:** To address our question comprehensively, we adopted a structured approach encompassing data management, analysis, and regression modeling.

**Findings:**

Sector Trends: Through ANOVA tests and sector-specific analyses, we identified significant variations in IPO prices across different sectors, shedding light on sectoral dynamics.

Top-Performing Companies: Teladoc emerged as a standout performer with the highest percentage price change over the five-year period, indicative of shifting market trends and consumer preferences.

Regression Insights: Regression analysis unveiled significant correlations between IPO-related variables and cumulative abnormal returns, providing insights into factors influencing post-IPO market performance.

## 2) Data Management

We used a dataset of Company IPOs Over Time from Kaggle, and we downloaded a corresponding dataset containing financial information of these companies from WRDS. We cleaned, manipulated and formatted both datasets before merging them according to each company's Ticker symbol, and we focused on the data from the year of the IPO to 5 years after the IPO year for each company.

## i) Clean ipo_kaggle_data

We established a new dataset, proj.ipo_kaggle_data_clean, with specific columns retained and ordered. We renamed columns for improved clarity and formatted financial metrics to enhance readability (e.g., IPO_Price in dollar format). To ensure consistency, we simplified market names, consolidating variations of NASDAQ and NYSE. We removed rows with missing State values to ensure that only companies in the US were kept. Additionally, we derived a new IPO_Year column from IPO_Date to facilitate time-based analysis. Lastly, we only retained necessary columns such as Company_Name, Sector, Industry, State, Ticker, Market, IPO_Price, IPO_Shares, IPO_Offer_Amount, IPO_Date and IPO_Year.

The processed dataset proj.ipo_kaggle_data_clean is the output of these operations. It contains cleaned and organized IPO data with renamed columns, formatted financial metrics, simplified market names, and rows with complete information. The dataset is sorted by IPO_Year, making it suitable for further analysis, such as IPO trends over time or sector-specific studies.

Total rows: 655 Total columns: 11 Rows 1-100

| | Company_Name | Sector | Industry | State | Ticker | Market | IPO_Price | IPO_Shares | IPO_Offer_Amount | IPO_Date | IPO_Year |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | GENERAC HOLDINGS INC. | Industrials | Diversified Industrials | WI | GNRC | NYSE | $13.00 | 18,750,000 | $243,750,000.00 | 2010-02-11 | 2010 |
| 2 | QUINSTREET, INC | Technology | Internet Content & Information | CA | QNST | NASDAQ | $15.00 | 10,000,000 | $150,000,000.00 | 2010-02-11 | 2010 |
| 3 | TERRENO REALTY CORP | Real Estate | REIT - Industrial | CA | TRNO | NYSE | $20.00 | 8,750,000 | $175,000,000.00 | 2010-02-10 | 2010 |
| 4 | PIEDMONT OFFICE REALTY TRUST, INC. | Real Estate | REIT - Office | GA | PDM | NYSE | $14.50 | 12,000,000 | $174,000,000.00 | 2010-02-10 | 2010 |
| 5 | IRONWOOD PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Specialty & Generic | MA | IRWD | NASDAQ | $11.25 | 16,666,667 | $187,500,004.00 | 2010-02-03 | 2010 |
| 6 | SS&C TECHNOLOGIES HOLDINGS INC | Technology | Software - Infrastructure | CT | SSNC | NASDAQ | $15.00 | 10,725,000 | $160,875,000.00 | 2010-03-31 | 2010 |
| 7 | FIRST INTERSTATE BANCSYSTEM INC | Financial Services | Banks - Regional - US | MT | FIBK | NASDAQ | $14.50 | 10,000,000 | $145,000,000.00 | 2010-03-24 | 2010 |
| 8 | CALIX, INC | Technology | Communication Equipment | CA | CALX | NYSE | $13.00 | 6,328,932 | $82,276,116.00 | 2010-03-24 | 2010 |
| 9 | MAXLINEAR INC | Technology | Semiconductors | CA | MXL | NYSE | $14.00 | 6,444,100 | $90,217,400.00 | 2010-03-24 | 2010 |
| 10 | AVEO PHARMACEUTICALS INC | Healthcare | Biotechnology | MA | AVEO | NASDAQ | $9.00 | 9,000,000 | $81,000,000.00 | 2010-03-12 | 2010 |
| 11 | ALPHA & OMEGA SEMICONDUCTOR LTD | Technology | Semiconductors | CA | AOSL | NASDAQ | $18.00 | 5,085,985 | $91,547,730.00 | 2010-04-29 | 2010 |
| 12 | CODEXIS INC | Healthcare | Biotechnology | CA | CDXS | NASDAQ | $13.00 | 6,000,000 | $78,000,000.00 | 2010-04-22 | 2010 |
| 13 | ALIMERA SCIENCES INC | Healthcare | Drug Manufacturers - Major | GA | ALIM | NASDAQ | $11.00 | 6,550,000 | $72,050,000.00 | 2010-04-22 | 2010 |
| 14 | SPS COMMERCE INC | Technology | Software - Application | MN | SPSC | NASDAQ | $12.00 | 4,096,694 | $49,160,328.00 | 2010-04-22 | 2010 |
| 15 | CHATHAM LODGING TRUST | Real Estate | REIT - Hotel & Motel | FL | CLDT | NYSE | $20.00 | 7,500,000 | $150,000,000.00 | 2010-04-16 | 2010 |
| 16 | PRIMERICA, INC. | Financial Services | Insurance - Life | GA | PRI | NYSE | $15.00 | 21,360,000 | $320,400,000.00 | 2010-04-01 | 2010 |
| 17 | GENMARK DIAGNOSTICS, INC. | Healthcare | Medical Devices | CA | GNMK | NASDAQ | $6.00 | 4,600,000 | $27,600,000.00 | 2010-05-28 | 2010 |
| 18 | EXPRESS, INC. | Consumer Cyclical | Apparel Stores | OH | EXPR | NYSE | $17.00 | 16,000,000 | $272,000,000.00 | 2010-05-13 | 2010 |
| 19 | ROADRUNNER TRANSPORTATION SYSTEMS, INC. | Industrials | Trucking | WI | RRTS | NYSE | $14.00 | 10,600,644 | $148,409,016.00 | 2010-05-13 | 2010 |
| 20 | TELENAV, INC. | Technology | Internet Content & Information | CA | TNAV | NASDAQ | $8.00 | 7,000,000 | $56,000,000.00 | 2010-05-13 | 2010 |
| 21 | DOUGLAS DYNAMICS, INC | Consumer Cyclical | Auto Parts | WI | PLOW | NYSE | $11.25 | 10,000,000 | $112,500,000.00 | 2010-05-05 | 2010 |
| 22 | TESLA, INC. | Consumer Cyclical | Auto Manufacturers | CA | TSLA | NASDAQ | $17.00 | 13,300,000 | $226,100,000.00 | 2010-06-29 | 2010 |
| 23 | OASIS PETROLEUM INC. | Energy | Oil & Gas E&P | TX | OAS | NYSE | $14.00 | 42,000,000 | $588,000,000.00 | 2010-06-17 | 2010 |
| 24 | GREEN DOT CORP | Financial Services | Credit Services | CA | GDOT | NYSE | $36.00 | 4,558,050 | $164,089,800.00 | 2010-07-22 | 2010 |
| 25 | AMERESCO, INC. | Industrials | Engineering & Construction | MA | AMRC | NYSE | $10.00 | 8,696,820 | $86,968,200.00 | 2010-07-22 | 2010 |
| 26 | WHITESTONE REIT | Real Estate | REIT - Retail | TX | WSR | NYSE | $12.00 | 2,200,000 | $26,400,000.00 | 2010-08-26 | 2010 |
| 27 | ELECTROMED, INC. | Healthcare | Medical Devices | MN | ELMD | NASDAQ | $4.00 | 1,700,000 | $6,800,000.00 | 2010-08-13 | 2010 |
| 28 | REALPAGE INC | Technology | Software - Application | TX | RP | NASDAQ | $11.00 | 12,300,000 | $135,300,000.00 | 2010-08-12 | 2010 |
| 29 | AMYRIS, INC. | Basic Materials | Specialty Chemicals | CA | AMRS | NASDAQ | $16.00 | 5,300,000 | $84,800,000.00 | 2010-09-28 | 2010 |
| 30 | CORESITE REALTY CORP | Real Estate | REIT - Office | CO | COR | NYSE | $16.00 | 16,900,000 | $270,400,000.00 | 2010-09-23 | 2010 |
| 31 | PACIFIC BIOSCIENCES OF CALIFORNIA, INC. | Healthcare | Diagnostics & Research | CA | PACB | NASDAQ | $16.00 | 12,500,000 | $200,000,000.00 | 2010-10-27 | 2010 |
| 32 | VERA BRADLEY, INC. | Consumer Cyclical | Footwear & Accessories | IN | VRA | NASDAQ | $16.00 | 11,000,000 | $176,000,000.00 | 2010-10-21 | 2010 |
| 33 | TOWER INTERNATIONAL, INC. | Consumer Cyclical | Auto Parts | MI | TOWR | NYSE | $13.00 | 6,250,000 | $81,250,000.00 | 2010-10-15 | 2010 |
| 34 | ELLINGTON FINANCIAL INC. | Financial Services | Specialty Finance | CT | EFC | NYSE | $22.50 | 4,500,000 | $101,250,000.00 | 2010-10-08 | 2010 |
| 35 | STANDARD AVB FINANCIAL CORP. | Financial Services | Banks - Regional - US | PA | STND | Nasdaq Smallcap Market | $10.00 | 3,450,000 | $34,500,000.00 | 2010-10-07 | 2010 |
| 36 | KEYW HOLDING CORP | Technology | Software - Application | MD | KEYW | NASDAQ | $10.00 | 9,100,000 | $91,000,000.00 | 2010-10-01 | 2010 |

## i) Clean stock_data

For the next step, we processed and analyzed stock data from 2008 to 2020. We cleaned and organized the data to produce a final dataset containing cumulative metrics.

We derived new columns from the date variable to extract Year, Month, and Day. We created an intermediate dataset, stock_intermediate, to calculate the abnormal return (abnret) as the difference between RET and vwretd. This dataset focuses on records with valid Tickers between 2008 and 2020. We then sorted the data by Ticker and Year.

We established a new dataset, yearly_averages, to compute yearly averages for metrics like avg_PRC (average price), avg_VOL (average volume), and avg_abnret (average abnormal return). To calculate these averages, we accumulated sums for each year and divided them by the count of records. We also formatted them accordingly for readability. In the final dataset,

proj.cum_stock, we calculated cumulative abnormal returns (Cum_AbnRet) by accumulating avg_abnret for each year. The cumulative sum is reset at the start of each Ticker.

The processed output is stored as proj.cum_stock. This dataset contains cumulative stock metrics, including cumulative abnormal returns, sorted by Year and Ticker. The cleaned dataset can be used for further analysis, such as assessing stock performance trends or comparing cumulative returns across different years and tickers.

Total rows: 5208 Total columns: 6

| | TICKER | Year | avg_PRC | avg_VOL | avg_abnret | Cum_AbnRet |
|---|---|---|---|---|---|---|
| 1 | AAOI | 2013 | 12.72 | 16570.50 | 0.08 | 0.0794 |
| 2 | AAOI | 2014 | 18.55 | 53414.58 | -0.00 | -0.0039 |
| 3 | AAOI | 2015 | 16.68 | 67772.75 | 0.05 | 0.0483 |
| 4 | AAOI | 2016 | 16.71 | 83939.00 | 0.03 | 0.0341 |
| 5 | AAOI | 2017 | 54.78 | 521182.3 | 0.07 | 0.0719 |
| 6 | AAOI | 2018 | 30.76 | 271117.6 | -0.04 | -0.0424 |
| 7 | AAOI | 2019 | 11.38 | 146671.4 | -0.03 | -0.0287 |
| 8 | AAOI | 2020 | 10.08 | 169332.3 | -0.03 | -0.0269 |
| 9 | AAT | 2011 | 21.05 | 64685.67 | 0.00 | 0.0014 |
| 10 | AAT | 2012 | 24.93 | 31776.42 | 0.02 | 0.0170 |
| 11 | AAT | 2013 | 31.37 | 41056.50 | -0.01 | -0.0090 |
| 12 | AAT | 2014 | 35.24 | 33520.25 | 0.01 | 0.0146 |
| 13 | AAT | 2015 | 40.70 | 40142.00 | 0.00 | 0.0016 |
| 14 | AAT | 2016 | 41.08 | 40440.42 | 0.00 | 0.0026 |
| 15 | AAT | 2017 | 40.62 | 50910.92 | -0.02 | -0.0230 |
| 16 | AAT | 2018 | 37.00 | 54157.75 | 0.01 | 0.0128 |
| 17 | AAT | 2019 | 46.11 | 71778.42 | -0.01 | -0.0085 |
| 18 | AAT | 2020 | 29.12 | 76715.00 | -0.04 | -0.0383 |
| 19 | ABAT | 2008 | 3.81 | 99424.36 | 0.03 | 0.0318 |
| 20 | ABAT | 2009 | 3.39 | 184814.6 | 0.02 | 0.0211 |
| 21 | ABAT | 2010 | 3.65 | 182113.2 | -0.02 | -0.0165 |
| 22 | ABAT | 2011 | . | 367349.8 | -0.09 | -0.0922 |
| 23 | ABC | 2008 | 39.16 | 356971.3 | 0.02 | 0.0218 |
| 24 | ABC | 2009 | 27.13 | 556718.8 | 0.01 | 0.0105 |
| 25 | ABC | 2010 | 30.32 | 764275.4 | 0.01 | 0.0105 |
| 26 | ABC | 2011 | 38.91 | 563254.3 | 0.01 | 0.0095 |
| 27 | ABC | 2012 | 39.28 | 545449.3 | 0.00 | 0.0019 |
| 28 | ABC | 2013 | 57.54 | 418287.5 | 0.02 | 0.0207 |
| 29 | ABC | 2014 | 75.82 | 357486.2 | 0.01 | 0.0147 |
| 30 | ABC | 2015 | 103.69 | 430192.2 | 0.01 | 0.0148 |
| 31 | ABC | 2016 | 81.80 | 534898.8 | -0.03 | -0.0292 |
| 32 | ABC | 2017 | 87.17 | 373652.4 | 0.00 | 0.0024 |
| 33 | ABC | 2018 | 87.86 | 324870.8 | -0.01 | -0.0084 |
| 34 | ABC | 2019 | 82.84 | 279044.8 | -0.01 | -0.0081 |
| 35 | ABC | 2020 | 94.60 | 271714.8 | -0.00 | -0.0048 |
| 36 | ABTX | 2015 | 23.70 | 14742.00 | -0.01 | -0.0109 |

## iii) Merge ipo_kaggle_data_clean and cum_stock

In the final step of data preparation we merged both datasets, creating a comprehensive dataset with related metrics.

We used PROC SQL to create a new table, and merged data by performing a left join on proj.ipo_kaggle_data_clean (the IPO dataset) and proj.cum_stock (the stock dataset). The join was based on the Ticker column and included stock data from the IPO year to five years after.

This merged dataset contains information from both the IPO and stock data, including company details, financial metrics, and cumulative stock returns.

The resulting dataset, proj.final_merged_data, integrates IPO information with stock metrics, providing a broader view of the financial landscape around IPOs. The data merge is designed to explore stock performance relative to IPO events, allowing for further analysis of trends and correlations.

Total rows: 1995 Total columns: 17 Rows 1-100

| | Company_Name | Sector | Industry | State | Ticker | Market | IPO_Price | IPO_Shares | IPO_Offer_Amount | IPO_Date | IPO_Year | Stock_Start_Year |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | APPLIED OPTOELECTRONICS, INC. | Technology | Semiconductors | TX | AAOI | NASDAQ | $10.00 | 3,600,000 | $36,000,000.00 | 2013-09-26 | 2013 | 2011 |
| 2 | APPLIED OPTOELECTRONICS, INC. | Technology | Semiconductors | TX | AAOI | NASDAQ | $10.00 | 3,600,000 | $36,000,000.00 | 2013-09-26 | 2013 | 2011 |
| 3 | APPLIED OPTOELECTRONICS, INC. | Technology | Semiconductors | TX | AAOI | NASDAQ | $10.00 | 3,600,000 | $36,000,000.00 | 2013-09-26 | 2013 | 2011 |
| 4 | AMERICAN ASSETS TRUST, INC. | Real Estate | REIT - Retail | CA | AAT | NYSE | $20.50 | 27,500,000 | $563,750,000.00 | 2011-01-13 | 2011 | 2009 |
| 5 | AMERICAN ASSETS TRUST, INC. | Real Estate | REIT - Retail | CA | AAT | NYSE | $20.50 | 27,500,000 | $563,750,000.00 | 2011-01-13 | 2011 | 2009 |
| 6 | AMERICAN ASSETS TRUST, INC. | Real Estate | REIT - Retail | CA | AAT | NYSE | $20.50 | 27,500,000 | $563,750,000.00 | 2011-01-13 | 2011 | 2009 |
| 7 | ALLEGIANCE BANCSHARES, INC. | Financial Services | Banks - Regional - US | TX | ABTX | NASDAQ | $21.00 | 2,600,000 | $54,600,000.00 | 2015-10-08 | 2015 | 2013 |
| 8 | ALLEGIANCE BANCSHARES, INC. | Financial Services | Banks - Regional - US | TX | ABTX | NASDAQ | $21.00 | 2,600,000 | $54,600,000.00 | 2015-10-08 | 2015 | 2013 |
| 9 | ALLEGIANCE BANCSHARES, INC. | Financial Services | Banks - Regional - US | TX | ABTX | NASDAQ | $21.00 | 2,600,000 | $54,600,000.00 | 2015-10-08 | 2015 | 2013 |
| 10 | ACACIA COMMUNICATIONS, INC. | Technology | Communication Equipment | MA | ACIA | NASDAQ | $23.00 | 4,500,000 | $103,500,000.00 | 2016-05-13 | 2016 | 2014 |
| 11 | ACACIA COMMUNICATIONS, INC. | Technology | Communication Equipment | MA | ACIA | NASDAQ | $23.00 | 4,500,000 | $103,500,000.00 | 2016-05-13 | 2016 | 2014 |
| 12 | ACACIA COMMUNICATIONS, INC. | Technology | Communication Equipment | MA | ACIA | NASDAQ | $23.00 | 4,500,000 | $103,500,000.00 | 2016-05-13 | 2016 | 2014 |
| 13 | ACM RESEARCH, INC. | Technology | Semiconductor Equipment & Materials | CA | ACMR | NASDAQ | $5.60 | 2,000,000 | $11,200,000.00 | 2017-11-03 | 2017 | 2015 |
| 14 | ACM RESEARCH, INC. | Technology | Semiconductor Equipment & Materials | CA | ACMR | NASDAQ | $5.60 | 2,000,000 | $11,200,000.00 | 2017-11-03 | 2017 | 2015 |
| 15 | ACM RESEARCH, INC. | Technology | Semiconductor Equipment & Materials | CA | ACMR | NASDAQ | $5.60 | 2,000,000 | $11,200,000.00 | 2017-11-03 | 2017 | 2015 |
| 16 | ACLARIS THERAPEUTICS, INC. | Healthcare | Biotechnology | PA | ACRS | NASDAQ | $11.00 | 5,000,000 | $55,000,000.00 | 2015-10-07 | 2015 | 2013 |
| 17 | ACLARIS THERAPEUTICS, INC. | Healthcare | Biotechnology | PA | ACRS | NASDAQ | $11.00 | 5,000,000 | $55,000,000.00 | 2015-10-07 | 2015 | 2013 |
| 18 | ACLARIS THERAPEUTICS, INC. | Healthcare | Biotechnology | PA | ACRS | NASDAQ | $11.00 | 5,000,000 | $55,000,000.00 | 2015-10-07 | 2015 | 2013 |
| 19 | ACELRX PHARMACEUTICALS INC | Healthcare | Medical Devices | CA | ACRX | NASDAQ | $5.00 | 8,000,000 | $40,000,000.00 | 2011-02-11 | 2011 | 2009 |
| 20 | ACELRX PHARMACEUTICALS INC | Healthcare | Medical Devices | CA | ACRX | NASDAQ | $5.00 | 8,000,000 | $40,000,000.00 | 2011-02-11 | 2011 | 2009 |
| 21 | ACELRX PHARMACEUTICALS INC | Healthcare | Medical Devices | CA | ACRX | NASDAQ | $5.00 | 8,000,000 | $40,000,000.00 | 2011-02-11 | 2011 | 2009 |
| 22 | ADAMAS PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Specialty & Generic | CA | ADMS | NASDAQ | $16.00 | 3,000,000 | $48,000,000.00 | 2014-04-10 | 2014 | 2012 |
| 23 | ADAMAS PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Specialty & Generic | CA | ADMS | NASDAQ | $16.00 | 3,000,000 | $48,000,000.00 | 2014-04-10 | 2014 | 2012 |
| 24 | ADAMAS PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Specialty & Generic | CA | ADMS | NASDAQ | $16.00 | 3,000,000 | $48,000,000.00 | 2014-04-10 | 2014 | 2012 |
| 25 | ADOMANI, INC. | Consumer Cyclical | Auto Parts | CA | ADOM | NASDAQ | $5.00 | 2,852,275 | $14,261,375.00 | 2017-06-15 | 2017 | 2015 |
| 26 | ADOMANI, INC. | Consumer Cyclical | Auto Parts | CA | ADOM | NASDAQ | $5.00 | 2,852,275 | $14,261,375.00 | 2017-06-15 | 2017 | 2015 |
| 27 | ADOMANI, INC. | Consumer Cyclical | Auto Parts | CA | ADOM | NASDAQ | $5.00 | 2,852,275 | $14,261,375.00 | 2017-06-15 | 2017 | 2015 |
| 28 | ADURO BIOTECH, INC. | Healthcare | Biotechnology | CA | ADRO | NASDAQ | $17.00 | 7,000,000 | $119,000,000.00 | 2015-04-15 | 2015 | 2013 |
| 29 | ADURO BIOTECH, INC. | Healthcare | Biotechnology | CA | ADRO | NASDAQ | $17.00 | 7,000,000 | $119,000,000.00 | 2015-04-15 | 2015 | 2013 |
| 30 | ADURO BIOTECH, INC. | Healthcare | Biotechnology | CA | ADRO | NASDAQ | $17.00 | 7,000,000 | $119,000,000.00 | 2015-04-15 | 2015 | 2013 |
| 31 | ADVANCED DISPOSAL SERVICES, INC. | Industrials | Waste Management | FL | ADSW | NYSE | $18.00 | 19,250,000 | $346,500,000.00 | 2016-10-06 | 2016 | 2014 |
| 32 | ADVANCED DISPOSAL SERVICES, INC. | Industrials | Waste Management | FL | ADSW | NYSE | $18.00 | 19,250,000 | $346,500,000.00 | 2016-10-06 | 2016 | 2014 |
| 33 | ADVANCED DISPOSAL SERVICES, INC. | Industrials | Waste Management | FL | ADSW | NYSE | $18.00 | 19,250,000 | $346,500,000.00 | 2016-10-06 | 2016 | 2014 |
| 34 | AERIE PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Major | NC | AERI | NASDAQ | $10.00 | 6,720,000 | $67,200,000.00 | 2013-10-25 | 2013 | 2011 |
| 35 | AERIE PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Major | NC | AERI | NASDAQ | $10.00 | 6,720,000 | $67,200,000.00 | 2013-10-25 | 2013 | 2011 |
| 36 | AERIE PHARMACEUTICALS INC | Healthcare | Drug Manufacturers - Major | NC | AERI | NASDAQ | $10.00 | 6,720,000 | $67,200,000.00 | 2013-10-25 | 2013 | 2011 |

**Explanation of variables in final_merged_data**

Company_Name: The name of the company.

Sector: The sector that the company belongs to.

Industry: The industry that the company belongs to.

State: The US state that the company is located in.

Ticker: The stock ticker symbol for the company.

Market: The market the company is listed on.

IPO_Price: The initial offering price of the company's shares at the time of the IPO.

IPO_Shares: The total number of shares offered in the IPO.

IPO_Offer_Amount: The total monetary value of the IPO

IPO_Date: The date when the IPO occurred.

IPO_Year: The year in which the IPO took place, derived from IPO_Date.

avg_PRC: The average stock price within the specified time frame.

avg_VOL: The average trading volume within the specified time frame.

avg_abnret: The average abnormal return; the deviation of the stock's return from an expected return.

Cum_AbnRet: The cumulative abnormal return.

Year: The year of each company's financial information, up to five years after the year of IPO for each company.

# 3) Data Analysis

We're examining companies that went public in 2015 to analyze sector trends. We're identifying the top companies in each sector with the highest average prices to understand sector performance. Then, we're evaluating the average stock price for these companies in 2020 to assess how their stock prices have changed over 5 years. Finally, we're determining how each sector has performed between 2015 and 2020 based on the performance of these companies and identifying the best-performing sector.

## i) Calculated Descriptive statistics

We use the PROC MEANS procedure to calculate summary statistics for variables from the merged_dataset. Specifically, we compute the count (N), mean, standard deviation (stddev), minimum, and maximum values for the variables IPO_Price, avg_PRC, avg_VOL, avg_abnret, and Cum_AbnRet. This analysis helps in summarizing the distribution and central tendency of the specified variables, providing insights into their overall characteristics within the dataset.

**The MEANS Procedure**

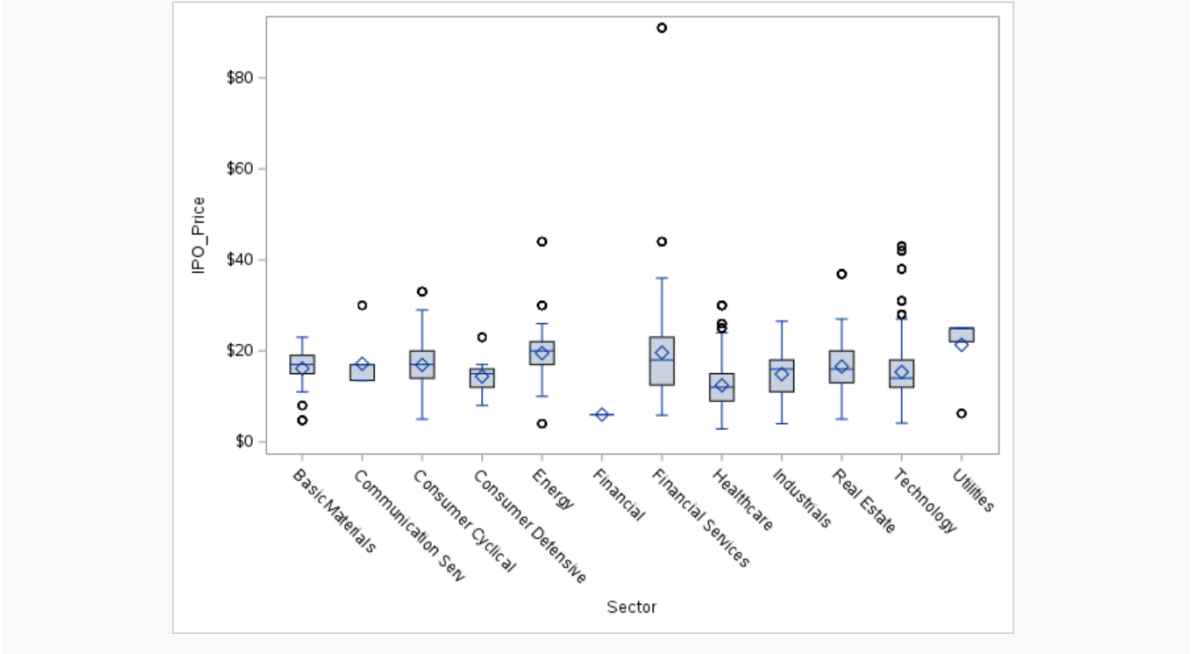
| Variable | N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|---|
| IPO_Price | 4704 | 15.5748278 | 6.8743584 | 2.8500000 | 91.0000000 |
| avg_PRC | 4558 | 25.8074552 | 33.8022817 | -0.3525000 | 713.2208383 |
| avg_VOL | 4691 | 221057.46 | 600975.48 | 317.5000000 | 12595815.25 |
| avg_abnret | 4704 | 0.0036777 | 0.0586334 | -0.2783756 | 1.2506945 |
| Cum_AbnRet | 4704 | 0.0036777 | 0.0586334 | -0.2783756 | 1.2506945 |

## ii) Performed ANOVA test

Performed ANOVA test and generated box plot to determine if there are statistically significant differences in IPO prices across different sectors or industries.

For the code segment, we employ the PROC SGPlot procedure to generate a vertical box plot (vbox) using the dataset. Specifically, we create a box plot of IPO_Price grouped by the variable Sector. This visualization aids in understanding the distribution of IPO prices across different sectors.

In the Anova test, it is evident that there is least price variation in communication services, consumer defensive & utilities sectors.




## iii) Top companies with highest average price at the time of IPO (Year = 2015)

For this step, we aim to identify the top company in each sector with the highest average price at the time of IPO for the year 2015.

We begin by filtering the dataset to exclusively encompass IPOs that took place in the year 2015. Subsequently, we calculate the average price (avg_PRC) for each company within every sector, utilizing the filtered dataset.

Following this, we sort the dataset by sector and average price in descending order to pinpoint the company with the highest average price in each sector. We employ the "if first.Sector" condition to single out the first record for each sector.We then present the resulting dataset, which delineates the top company in each sector alongside its average price.

Subsequently, to ensure data integrity and avoid redundancy, we eliminate the initial row of data for each sector. Furthermore, we update the company name for the "Consumer Defensive" sector to "OLLIE'S BARGAIN OUTLET HOLDINGS, INC" in accordance with the provided code.
This process enables us to identify and present the top company in each sector with the highest average price at the time of IPO for the year 2015.

| Obs | Sector | Company_Name | Avg_PRC |
|---|---|---|---|
| 1 | Basic Materials | ENVIVA PARTNERS, LP | 28.03 |
| 2 | Consumer Cyclical | WINGSTOP INC. | 57.53 |
| 3 | Consumer Defensive | OLLIE'S BARGAIN OUTLET HOLDINGS, INC | 52.37 |
| 4 | Energy | GREEN PLAINS PARTNERS LP | 14.98 |
| 5 | Financial Services | HOULIHAN LOKEY, INC. | 39.20 |
| 6 | Healthcare | PENUMBRA INC | 111.06 |
| 7 | Industrials | TRANSUNION | 31.55 |
| 8 | Real Estate | COMMUNITY HEALTHCARE TRUST INC | 29.56 |
| 9 | Technology | TELADOC HEALTH, INC. | 62.22 |

## iv) Average price of stocks for same companies in the year 2020

In this step, we analyze the price of these company stocks after 5 years, specifically in 2020.

First, we extract data from the dataset, filtering for records from the year 2020. Additionally, we narrow down the selection to include only specific companies using their tickers (Ticker) such as 'PEN', 'EVA', 'WING', 'GPRE', 'HLI', 'TRU', 'CHCT', 'TDOC', and 'OLLI'. We retain the variables Sector, Company_Name, and Avg_PRC for analysis. Next, we present the resulting dataset, using the PROC PRINT procedure. This dataset contains information about the top companies in each sector with the highest average price in the year 2020.

This analysis provides insights into the performance of selected companies' stock prices five years after their initial public offering (IPO) and facilitates the identification of trends or standout performers in the market.

**Top Company in Each Sector with Highest Average Price (Year 2020)**

| Obs | Sector | Company_Name | avg_PRC |
|---|---|---|---|
| 1 | Real Estate | COMMUNITY HEALTHCARE TRUST INC | 43.78 |
| 2 | Basic Materials | ENVIVA PARTNERS, LP | 38.42 |
| 3 | Financial Services | HOULIHAN LOKEY, INC. | 58.15 |
| 4 | Consumer Defensive | OLLIE'S BARGAIN OUTLET HOLDINGS, INC. | 79.35 |
| 5 | Healthcare | PENUMBRA INC | 192.88 |
| 6 | Technology | TELADOC HEALTH, INC. | 181.58 |
| 7 | Industrials | TRANSUNION | 85.78 |
| 8 | Consumer Cyclical | WINGSTOP INC. | 122.30 |

## v) Top company and sector with highest percentage price change over 5 years (2015-2020)

In the final step, we aim to identify the top company and sector with the highest percentage price change over a span of 5 years, from 2015 to 2020.

Initially, we merge the datasets based on Company_Name and Sector using PROC SQL, creating a new dataset named price_comparison. This dataset combines company price information from 2015 and 2020. Next, we calculate the percentage change for each company's stock price over the 5-year period. This involves computing the percentage change between the prices in 2015 and 2020. We

handle cases where the Price_2015 is zero to prevent division by zero and assign a missing value (.) in such instances.

After computing the percentage change, we sort the data to identify the companies with the highest percentage change. Finally, we present the resulting dataset using PROC PRINT, showcasing the top companies along with their respective sectors, prices in 2015 and 2020, and the calculated percentage price change over the specified period.

This analysis provides insights into significant price fluctuations over the 5-year period, aiding in the identification of potential market trends and investment opportunities.

**Percentage Price Change from 2015 to 2020 for Top Companies**

| Obs | Company_Name | Sector | Price_2015 | Price_2020 | Percent_Change |
|---|---|---|---|---|---|
| 1 | TELADOC HEALTH, INC. | Technology | 62.22 | 181.58 | 191.832 |
| 2 | TRANSUNION | Industrials | 31.55 | 85.78 | 171.881 |
| 3 | WINGSTOP INC. | Consumer Cyclical | 57.53 | 122.30 | 112.592 |
| 4 | PENUMBRA INC | Healthcare | 111.06 | 192.88 | 73.678 |
| 5 | HOULIHAN LOKEY, INC. | Financial Services | 39.20 | 58.15 | 48.360 |
| 6 | COMMUNITY HEALTHCARE TRUST INC | Real Estate | 29.56 | 43.78 | 48.129 |
| 7 | ENVIVA PARTNERS, LP | Basic Materials | 28.03 | 38.42 | 37.051 |

Compared stock prices of top performing companies in 2015 to their prices in 2020. We see that Teladoc has the highest percentage change compared to the other top performing IPOs of 2015. This may be due to Teladoc's rise in popularity during the pandemic where people could not visit doctors in person.

In conclusion, by comparing average stock prices in 2015 and 2020, we identified the top companies in each sector and their performance over five years. Teladoc demonstrated the highest percentage price change, indicating significant growth, likely due to increased demand for virtual healthcare during the pandemic. These results highlight sector-specific trends, with notable growth in particular industries and companies, guiding our understanding of the best-performing sectors and investment opportunities over the five-year period.

# 4)Regression Analysis

## i) Data Selection

Firstly, we defined a vector named variables that selects a subset of relevant variables for our regression model. These chosen variables relate to Initial Public Offerings (IPOs), encompassing state, IPO price, number of IPO shares (IPO_Shares) , IPO issue amount (IPO_Offer_Amount), average price (avg_PRC), average trading volume (avg_VOL), and cumulative abnormal returns (Cum_AbnRet). This selection includes key data points that we often analyze in financial and investment scenarios to gain insights into IPO performance, market trends, and broader economic outcomes.
Our goal is to lay the groundwork for building a regression model that can predict or explain specific outcomes based on these selected variables.

```
# read data
IPO.df <- read.csv("final.csv")
head(IPO.df)
```

```
# Load the required libraries
library(dplyr)
# Select relevant variables for the regression model
variables<-c("State","IPO_Price","IPO_Shares","IPO_Offer_Amount","avg_PRC","avg_VOL","Cum_A
bnRe")
```

## ii) Data Processing

In this step, we processed the data by handling missing values, converting categorical variables, and transforming text-based financial metrics into numeric format.
First, we removed rows with missing values using na.omit(), ensuring only complete data for analysis. Next, the State variable was converted into a factor with as.factor(data$State), categorizing it for later use.
To prepare for quantitative analysis, we used gsub() to remove non-numeric characters from IPO_Price, IPO_Shares, and IPO_Offer_Amount, converting them into numeric values with as.numeric().

```
# Filter out any missing values
data <- na.omit(IPO.df[, variables])
sum(is.na(data))
str(data)
data$State <- as.factor(data$State)
# Convert chr to num
data$IPO_Price <- as.numeric(gsub("[^0-9.]", "", data$IPO_Price))
data$IPO_Shares <- as.numeric(gsub("[^0-9]", "", data$IPO_Shares))
data$IPO_Offer_Amount <- as.numeric(gsub("[^0-9.]", "", data$IPO_Offer_Amount))
```

## iii) Regression Model

Finally, the function is used to build a linear regression model to observe the effect of different variables on cumulative abnormal returns (Cum_abnret).
We constructed a linear regression model to investigate the effects of various variables on cumulative abnormal returns (Cum_AbnRet).
Our model <- lm(Cum_AbnRet ~ ., data = data) establishes a linear relationship between the dependent variable, Cum_AbnRet, and all other variables in the dataset, indicated by the period (~ .). By including all variables, we aim to explore which factors significantly impact cumulative abnormal returns.
This regression model helps us assess the strength and direction of the relationships between Cum_AbnRet and our predictors. Analyzing this model allows us to identify which variables play a crucial role in influencing cumulative abnormal returns.

## iv) Regression Analysis

Use function summary() to get the result of the regression model as below:

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.732e-01  4.101e-02   4.224 2.52e-05 ***
StateAZ          -1.776e-01  4.218e-02  -4.211 2.66e-05 ***
StateCA          -1.882e-01  3.925e-02  -4.796 1.74e-06 ***
StateCO          -2.083e-01  4.019e-02  -5.182 2.42e-07 ***
StateCT          -1.778e-01  4.131e-02  -4.305 1.76e-05 ***
StateDC          -1.741e-01  6.412e-02  -2.715 0.006681 **
StateFL          -1.406e-01  4.096e-02  -3.432 0.000613 ***
StateGA          -1.272e-01  4.120e-02  -3.088 0.002043 **
StateHI          -1.608e-01  6.365e-02  -2.527 0.011576 *
StateIA          -1.618e-01  4.923e-02  -3.286 0.001033 **
StateID          -2.118e-01  5.284e-02  -4.008 6.36e-05 ***
StateIL          -1.835e-01  4.053e-02  -4.528 6.31e-06 ***
StateIN          -1.621e-01  4.470e-02  -3.627 0.000294 ***
StateKS          -1.472e-01  4.501e-02  -3.270 0.001094 **
StateKY          -9.495e-02  4.881e-02  -1.945 0.051881 .
StateLA          -1.666e-01  6.414e-02  -2.598 0.009453 **
StateMA          -1.906e-01  3.948e-02  -4.827 1.49e-06 ***
StateMD          -1.863e-01  4.161e-02  -4.476 8.05e-06 ***
StateMI          -1.938e-01  4.373e-02  -4.431 9.93e-06 ***
StateMN          -1.759e-01  4.378e-02  -4.018 6.11e-05 ***
StateMO          -2.141e-01  5.297e-02  -4.042 5.51e-05 ***
StateMT          -1.908e-01  6.411e-02  -2.976 0.002958 **
StateNC          -1.596e-01  4.146e-02  -3.848 0.000123 ***
StateNE          -1.429e-01  6.410e-02  -2.230 0.025866 *
StateNH          -1.801e-01  5.286e-02  -3.406 0.000672 ***
StateNJ          -1.905e-01  4.051e-02  -4.702 2.76e-06 ***
StateNV          -1.885e-01  4.680e-02  -4.028 5.86e-05 ***
StateNY          -2.057e-01  3.960e-02  -5.194 2.28e-07 ***
StateOH          -1.687e-01  4.366e-02  -3.864 0.000115 ***
StateOK          -1.834e-01  4.178e-02  -4.390 1.20e-05 ***
StateOR          -1.774e-01  6.411e-02  -2.768 0.005701 **
StatePA          -1.898e-01  4.048e-02  -4.689 2.93e-06 ***
StateRI          -1.477e-01  6.536e-02  -2.260 0.023911 *
StateSC          -2.064e-01  5.265e-02  -3.920 9.17e-05 ***
StateSD          -1.290e-01  6.398e-02  -2.015 0.044010 *
StateTN          -1.492e-01  4.133e-02  -3.610 0.000314 ***
StateTX          -1.877e-01  3.921e-02  -4.788 1.82e-06 ***
StateUT          -2.024e-01  4.319e-02  -4.687 2.96e-06 ***
StateVA          -1.799e-01  4.325e-02  -4.160 3.32e-05 ***
StateWA          -1.550e-01  4.190e-02  -3.699 0.000222 ***
StateWI          -1.699e-01  4.170e-02  -4.074 4.82e-05 ***
IPO_Price        -3.293e-03  3.948e-04  -8.341  < 2e-16 ***
IPO_Shares        2.450e-10  2.509e-10   0.976 0.329038
IPO_Offer_Amount -1.949e-11  8.042e-12  -2.424 0.015446 *
avg_PRC           3.278e-03  1.515e-04  21.635  < 2e-16 ***
avg_VOL           1.526e-08  5.388e-09   2.833 0.004662 **
```

The p value smaller than the significance level means the coefficient of the variable is statistically significant. In other words, this explanatory variable has a significant impact on the dependent variable. The “***” means p<0.001, the “**” means p<0.01 and the “*” means p<0.05 and no stars means this coefficient is not statistically significant.

The result shows that most of the state variables have significant impact besides the state KY. The variables IPO_Price, IPO_Offer_Amount, avg_PRC and avg_VOL are significant. However, the IPO_Shares is not significant for the p-value is larger than 0.05.

The value of adjusted R squared is 0.2242 which means this regression model has poor fitting ability. Besides, the p-value for F-statistic is very small and lower than 0.001, so this model is still significant and reasonable.

```
Residual standard error: 0.08825 on 1920 degrees of freedom
Multiple R-squared:  0.2419,    Adjusted R-squared:  0.2242
F-statistic: 13.62 on 45 and 1920 DF,  p-value: < 2.2e-16
```

The max residual is relatively large which might indicate that there are some outliers. The low value of adjusted R squared might be caused by the non-linear relationship between some of the variables.

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.29445 -0.03911 -0.00217  0.03210  1.24539
```

In summary, our linear regression model reveals significant correlations between specific variables and cumulative abnormal returns (Cum_AbnRet).

We found that State, IPO_Price, and IPO_Offer_Amount had significant negative correlations with cumulative abnormal returns. This suggests that certain states, along with higher IPO prices and offer amounts, are associated with lower returns, indicating that these factors might impact performance negatively.

On the other hand, avg_PRC and avg_VOL showed significant positive correlations with cumulative abnormal returns. This implies that higher average prices and trading volumes tend to be linked with increased returns, pointing toward stronger market trends or greater investor engagement.

# 5) Summary

Our report delves into the intricacies of Initial Public Offerings (IPOs) and their profound impact on market performance. By meticulously managing, analyzing, and modeling data, we uncovered significant insights into sector trends, top-performing companies, and the relationships between IPO-related variables and cumulative abnormal returns. Notably, Teladoc emerged as a standout performer, reflecting the evolving market landscape. Our findings underscore the importance of understanding IPO dynamics for informed decision-making in the dynamic realm of finance.