**VERTEX PHARAMACEUTICALS – AUDIT QUALIFICATION ANALYSIS**

Done By

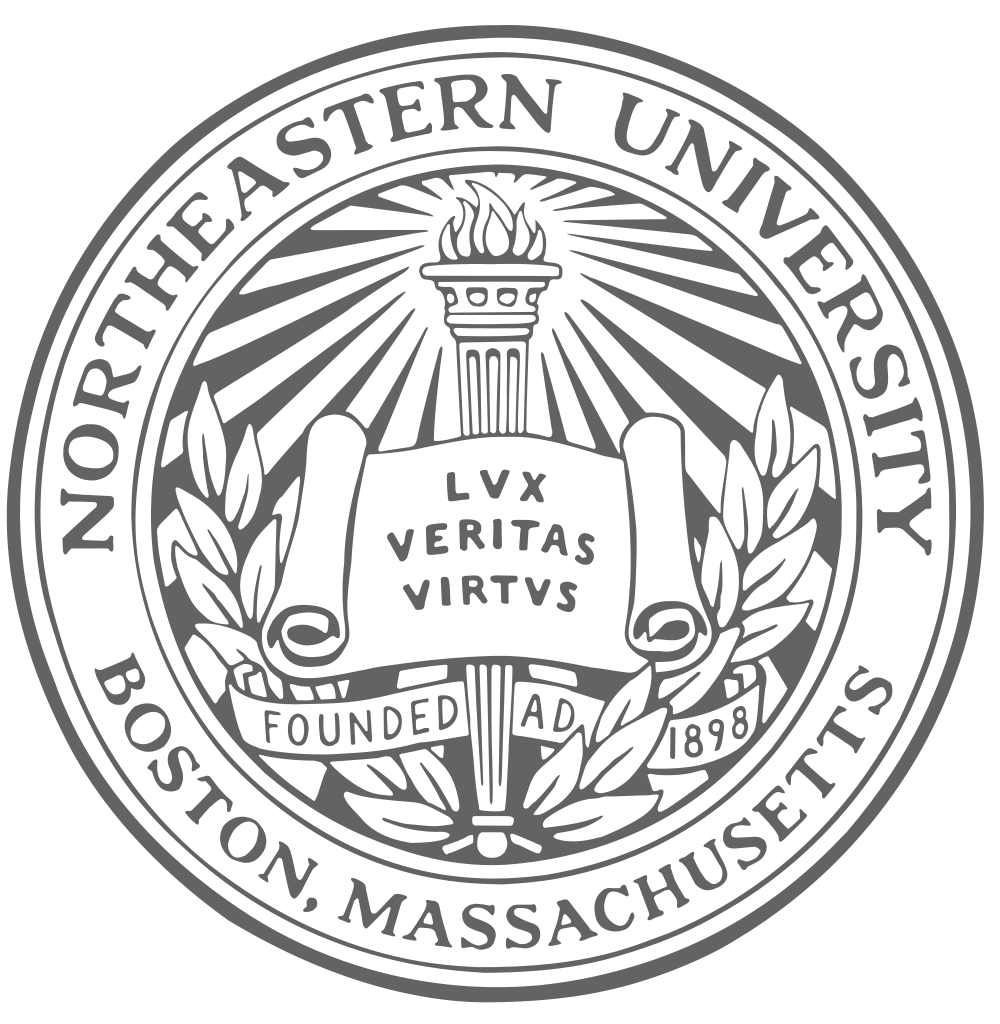
ANISH NITIN SOMAIAH

**WEEK 11 : FINAL PROJECT REPORT**

ALY 6080: 81190 Integrated Experiential Learning

Instructor: Mr. AMIN KARIMPOUR

Date: June 25, 2019



**INTRODUCTION:**

Vertex Pharmaceuticals, Inc. ,an American biopharmaceutical company located in Boston, Massachusetts, is one of the first biotech firms to use an explicit strategy of rational drug design rather than combinatorial chemistry. Being recognized as one of the greatest places to work in this industry, the contributions of this organization to the industry and in curation of genetic diseases has been very progressive. Vertex is known to have designed and developed the first medicines to treat the underlying cause of a rare significant life-threatening genetic disease named **CYSTIC FIBROSIS**. As per investment data the organization has earned an investment of 166.18M$ during the first quarter of 2019 from NASDAQ exchange for its upcoming projects in Research & Development.

From the company’s stock growth and from its claims we know that Vertex invests upto 72% of its stocks and investments into R&D as depicted below:



**PROJECT OBJECTIVE:**

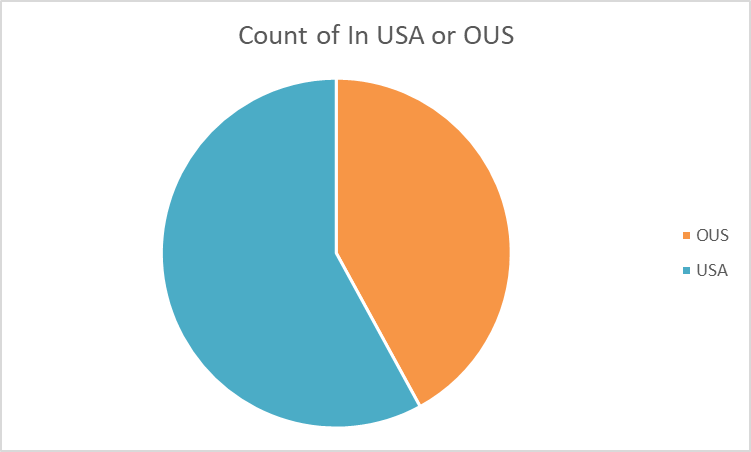
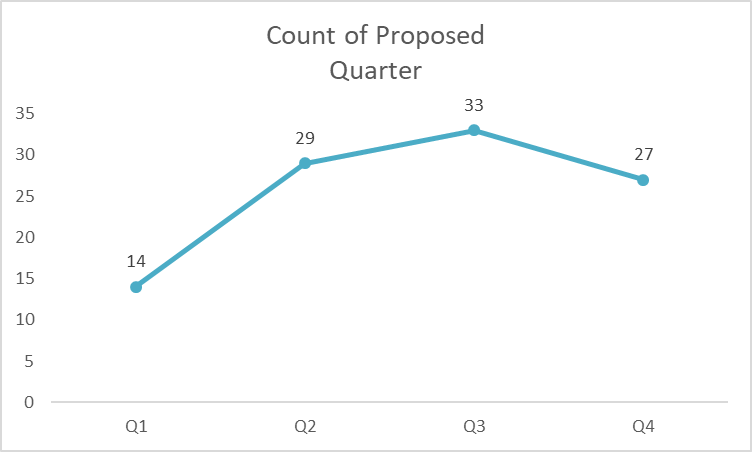
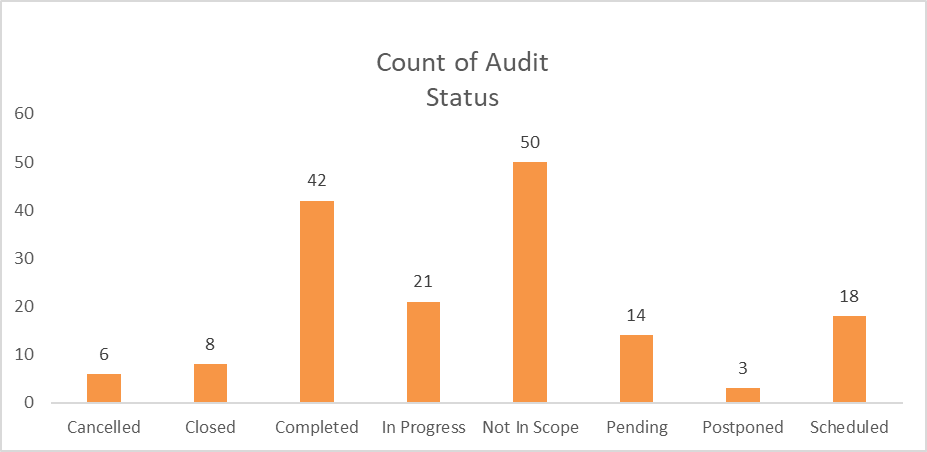
As per the companies claims, only 28% of its revenues are spent on administrative purposes. Now the organization has different vendors to procure the organization with materials, chemicals and general-purpose office supplies for the organization. Our objective is to **figure** **optimum solution to regulate the vendor activities for different scenarios, optimize the 28% of the revenue for efficient vendor performances, provide cost efficient methods to ensure sustained vendor supply activity.**

**PROJECT INPUT:**

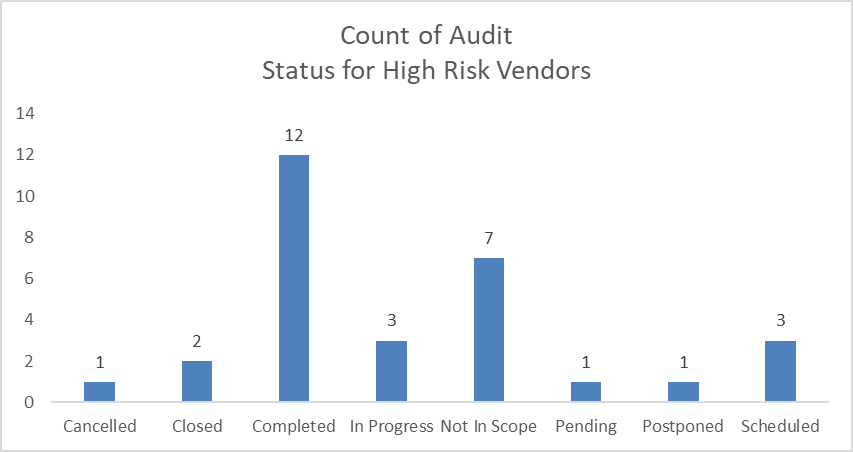
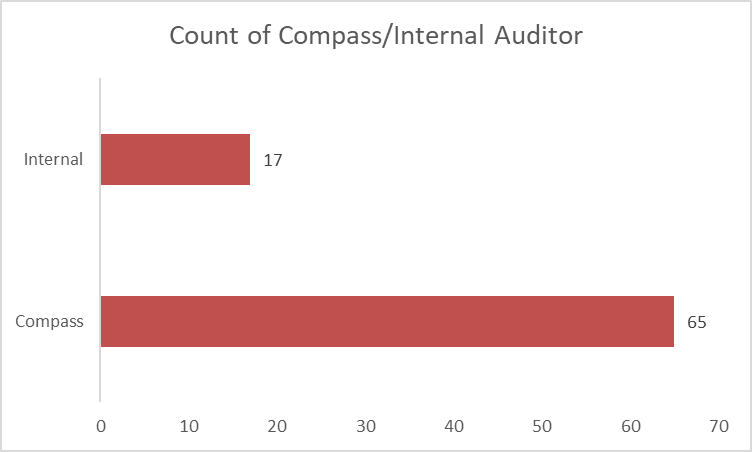
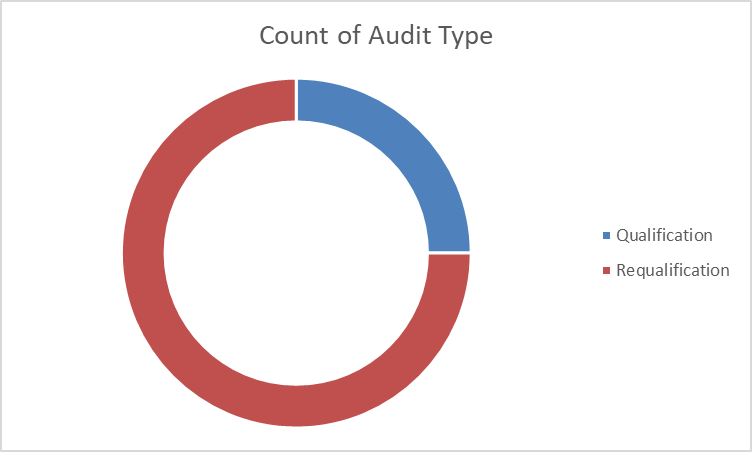
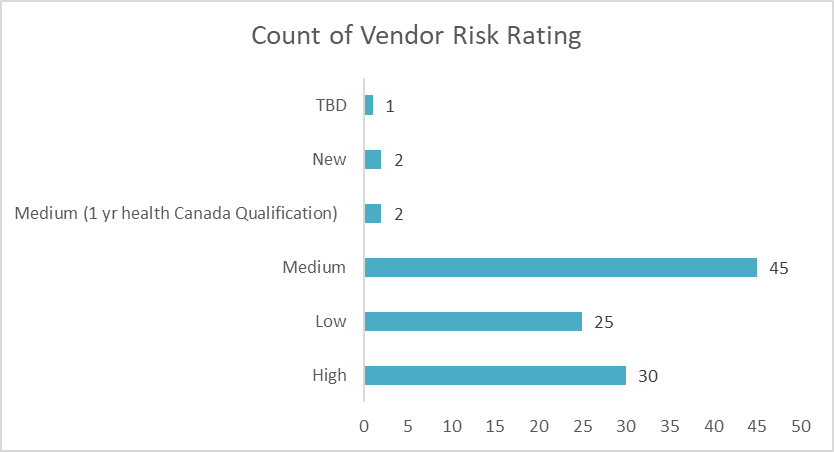
We are given 2 supply chain audit trackers one for GxP audits and the other for CSQA audits along with metrics such as Audit Status, In USA or OUS, GxP Area, Vendor Risk Rating, Audit Type, CSQA Audit Method, Proposed Quarter, Compass/Internal Auditor etc. and so on for us to analyze and provide metrics based on their behavioral patterns from 2015-2018.

**ANALYTICS STAGE 1 : STATISTICS OF 2018**

*Dashboard 1 : Visualization for Market, Quarter and Status*



*Dashboard 2 : Visualization of Vendor Risk, High Risk Status, Audit type and Auditor organization*



**SOLUTIONS FROM STAGE 1 ANALYSIS:**

* Around 58% of the audits are within USA and 42% are Outside USA in 2018.
* Q3 has the greatest number of audits for 2018 but the distribution of no.of audits is almost uniform throughout the year.
* Only 42/163 ( 25.7% ) audits have been completed. And 50/163 have been pushed out of scope. Our data doesn’t give any clarity about the reason about being declared not in scope. It might be because the products received from the vendors might be of low risk or not required for current R&D.
* There were a significant number of medium and high-risk vendors in 2018, with 29% high risk audits. So, audit status was drilled down for the high-risk vendors and was observed that 40% of the audits were completed which is a good sign. Hence concentration can be shifted towards the medium and low risked vendors.
* We also observe 75% of the audits were Requalification. This can be assumed that either the qualification audit was not implemented right or the product was essential for further R&D. This will be examined in phase 2 review where case wise trends are analyzed based on the statuses of the Requalification audits.
* We also observe that 65/82 audits in 2018 were done by a third-party company. If manpower is efficient within vertex organization, then a shift can be done between the manpower and the distribution of audit organization can be made uniform. This can save time in completing audits and cost cut that the company has to invest on third party organizations.

**ANALYTICS STAGE 2 : EXPLORATORY DATA ANALYTICS**

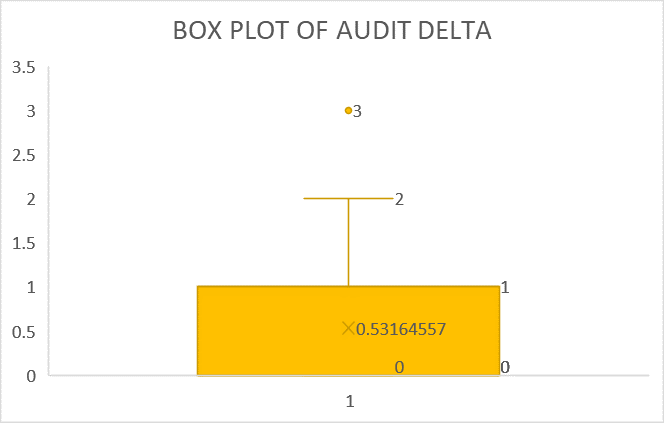
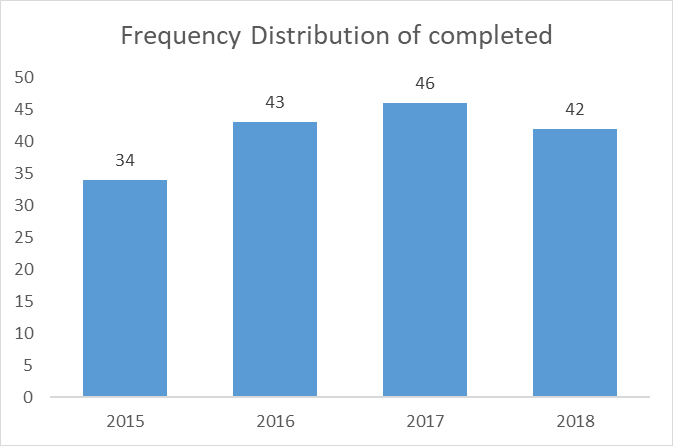
*Figure 1: Audit Status Count Distribution.*

From the above graph we infer that

* Out of the total 506 audits carried out since 2015, 165 is in completed status which is the highest of the lot. Hence we consider the completed status EDA to find out metrics to obtain the completion of other audits.
* We have 89 audits which became “Not in Scope”. This might be due to irrelevancy of supplier product or less demand at that particular time which is not clearly mentioned in the dataset. We also have 50 audits which were not done and 42 cancelled for which no clarity information is provided to us.

**39 audits in progress can happen as it is, and the metrics used in EDA of completed can be adopted to these cases along with the 22 pending and 20 scheduled to complete audit on time.**

*Figure 2: Histogram of audits year-wise Figure 3 : Box Plot of Audit Delta*



**From the graph we infer that the distribution is almost normal with peak no.of audits completed in 2017. Hence any standardized machine learning algorithm can be applied to this dataset for further predictions.**

**From the above box plot we infer that the average time between audit start and audit end date is 0.5-1 days with peak of 2 days. There was only one outlier among the completed audits where there was a gap of 3 days between audit start and end date.**

**NOTE: this is just time difference between audit start and audit end date and not the difference between actual intake and completion dates. This metric was considered to give a generalized time duration of any audit whether questionnaire or on-site.**

Now let us look at the individual Exploratory data analytics of questionnaire, on-site and total audit completion based on the dates given which are as follows:

* Questionnaire Delta = difference between Q sent and Q received
* Audit Delta = difference between Audit start date and Audit end date
* Audit Completion Delta = difference between date intake and date completion

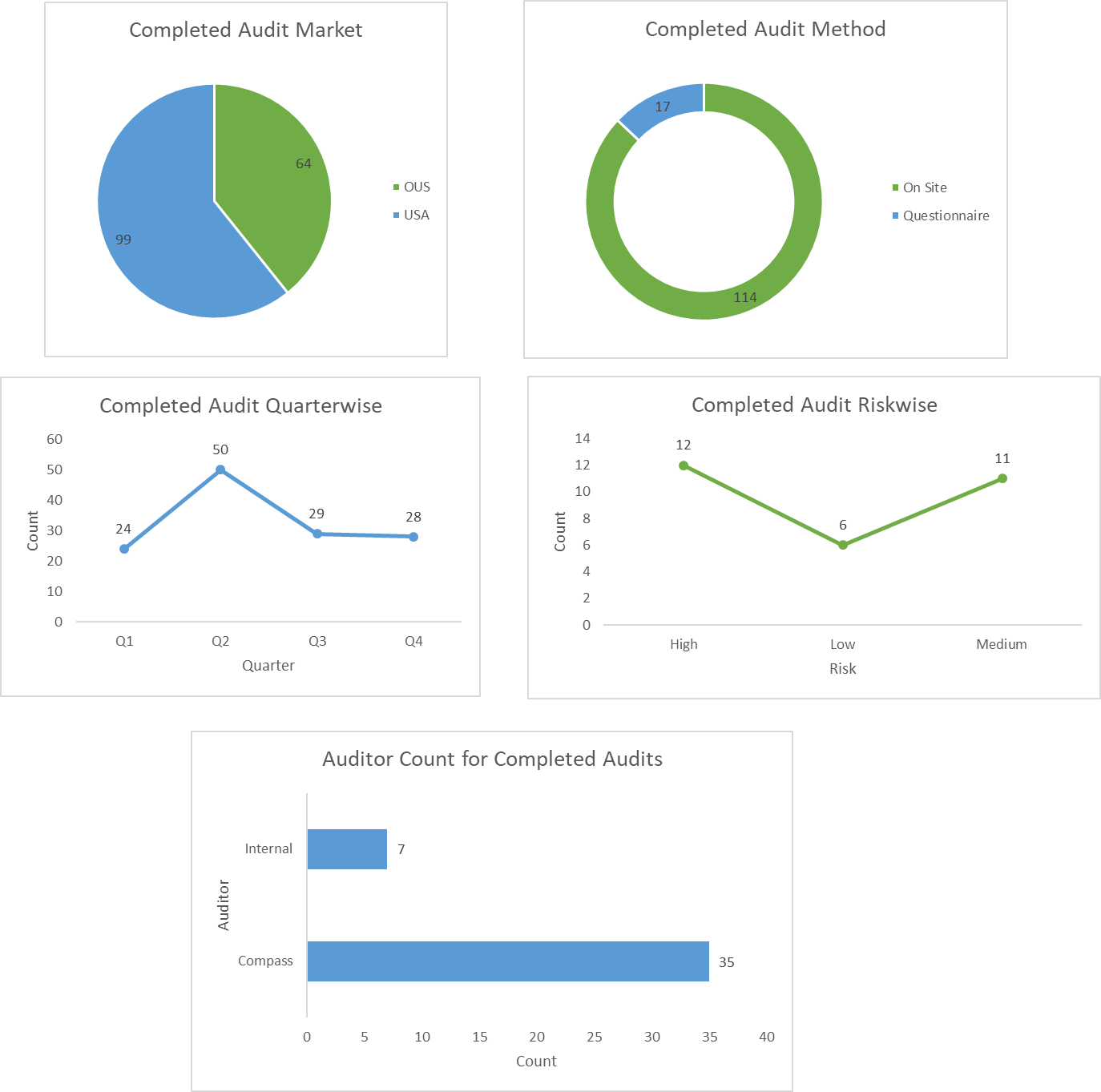
*Table 1 : EDA for completion based on dates*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *questionnaire delta* | |  | *audit delta* | |  | *audit completion delta* | |
| Mean | 2.40 |  | Mean | 0.532 |  | Mean | 103.46 |
| Standard Error | 1.01 |  | Standard Error | 0.054 |  | Standard Error | 9.50 |
| Median | 0.00 |  | Median | 0.000 |  | Median | 85.00 |
| Mode | 0.00 |  | Mode | 0.000 |  | Mode | 0.00 |
| Standard Deviation | 12.78 |  | Standard Deviation | 0.674 |  | Standard Deviation | 110.84 |
| Sample Variance | 163.42 |  | Sample Variance | 0.454 |  | Sample Variance | 12285.39 |
| Kurtosis | 49.40 |  | Kurtosis | 0.381 |  | Kurtosis | -0.59 |
| Skewness | 6.79 |  | Skewness | 1.021 |  | Skewness | 0.78 |
| Range | 107.00 |  | Range | 3.000 |  | Range | 369.00 |
| Minimum | 0.00 |  | Minimum | 0.000 |  | Minimum | 0.00 |
| Maximum | 107.00 |  | Maximum | 3.000 |  | Maximum | 369.00 |
| Sum | 387.00 |  | Sum | 84.000 |  | Sum | 14071.00 |
| Count | 161.00 |  | Count | 158.000 |  | Count | 136.00 |

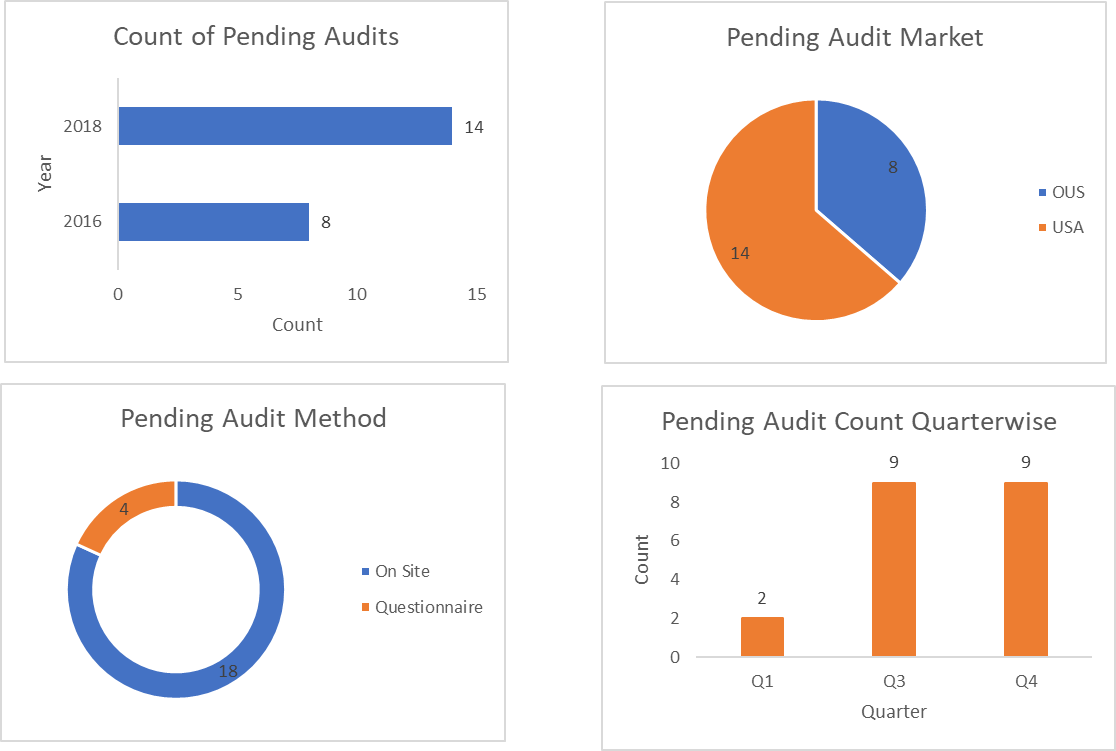
**SOLUTIONS FROM STAGE 2 ANALYSIS**

* The average time taken to finish a questionnaire is almost 2.4 days , for an audit is 0.5 days.
* But as close as it seems, based on the given data a trend has been observed that on an average an audit takes 103 days on an average to complete an audit from start to end. With a standard deviation of 111 meaning the dataset is widespread. This huge time difference is a significant challenge and the limitation is that we do not have any data to rectify this time gap.
* The distribution is also right skewed meaning audits are completed on a large scale at initial stage or initial quarters and a delay is observed in the later part of the year. The negative kurtosis indicates a less presence of outliers or smaller tail in the distribution.
* The main parameter to be considered is the standard error. If this is reduced from 9.5 to 0, we have a higher chance of shifting the distribution from right skew to normal.
* The maximum time taken to complete an audit is 369 days close to 1 year. This also has to be worked onto as the audit happened in 2017 which had many audits completed and above all the audit took place in USA.

**ANALYTICS STAGE 3 : TREND ANALYTICS BASED ON COMPLETED EDA**

*Figure 4: Dashboard 1 Complete Audit status*

*Figure 5: Dashboard 2 Pending Audit status*



**SOLUTIONS FROM STAGE 3 ANALYSIS**

* The USA vendors are more than that of the outside USA, hence transport costs within USA only should be considered for optimization of transport cost. Added to it are the major number of on-site audits hence transport cost optimization within USA are to be considered.
* We observe that High and Medium Risk Audits are completed more than low which proves a higher efficiency of audits taken place.
* We observe from the dashboard that most of the audits are completed on quarter 2 and there is a steep decline from Q2. This also proves the right skewed pattern of the completion distribution. Hence Q3 and Q4 cases are to be worked out to reach the metrics of Q2.
* We also observe that more no.of audits are completed by 3rd party organization which is actually pulling up the cost of audit qualification. Hence manpower of the internal auditors are to be distributed evenly to ensure completion of audits within budget.
* We observe from the above dashboard that pending status follows the same pattern as that of the completed status with more in USA, more on-site audits and Q2 completed and pending cases in 2016 and 2018 proving our initial frequency histogram.

**ANALYTICS STAGE 4 : AUDITOR PERFORMANCE ANALYTICS (GxP)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Audit Method | Count |  | Audit Status | Count |
| On Site | 83 |  | Cancelled | 5 |
| Questionnaire | 16 |  | Closed | 20 |
| Remote | 1 |  | Completed | 61 |
| Method Unknown | 7 |  | In Progress | 5 |
|  |  |  | Not in Scope | 7 |
|  |  |  | Pending | 2 |
|  |  |  | Scheduled | 7 |

We Observe that Auditor TB has done the maximum no.of GxP audits. His Metrics are as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Audit Status** | **In USA or OUS** | **audit completion delta** |
| 2018 | Pending | OUS |  |
| 2017 | Not in Scope | OUS |  |
| 2017 | Not in Scope | OUS |  |
| 2017 | Completed | OUS | 112 |
| 2017 | Completed | OUS | 147 |
| 2017 | Completed | OUS | 357 |
| 2017 | Completed | OUS | 154 |
| 2017 | Completed | OUS | 155 |

**SOLUTIONS FROM STAGE 4 ANALYSIS:**

* 107 audits done by 47 auditors who have done the GxP audits.
* 57 audits took place in USA and 50 outside of USA.
* From observation we find that it takes an average of 178 days to complete an onsite audit OUS and but 190 days to complete an audit within USA.
* Only 3/107 audits were carried out by Internal Auditors.
* From the above table we observe that TB has completed 5/8 audits which are Outside USA onsite audits and he has taken an average of 186 days to complete.
* He has done maximum no.of audits within the average timeframe of a GxP audit. Hence to optimize manpower as a majority are Compass auditors, 5 auditors who have done single audits and who has taken more than average no.of days to complete the audit can be assigned to a single auditor.
* If this metric is used, we have 14 audits which are pending which are assigned to 14 different auditors, these audits can be done by approximately 3 auditors who have already completed their audits and have nothing scheduled for them and these 14 auditors can be moved to the 200 remaining audits where we do not have auditor details or CSQA audits to improve efficiency with same cost or these auditors can be removed to cut cost for vertex.

**ANALYTICS STAGE 5 : AUDITOR PERFORMANCE ANALYTICS (CSQA)**

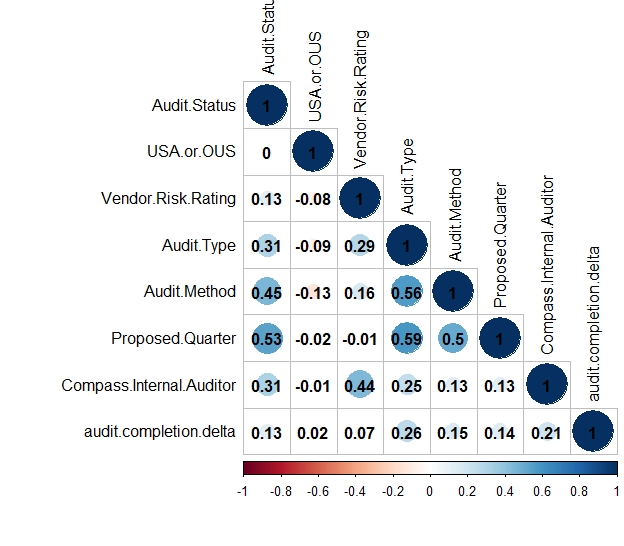
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Audit Type | Count |  | Audit Status | Count |
| Internal | 5 |  | Cancelled | 7 |
| Qualification | 19 |  | Closed | 34 |
| Requalification | 54 |  | Completed | 23 |
|  |  |  | In Progress | 7 |
|  |  |  | Not In Scope | 1 |
| Audit Method | Count |  | On Hold | 2 |
| On Site | 54 |  | Pending | 1 |
| Questionnaire | 24 |  | Scheduled | 3 |

**SOLUTIONS FROM STAGE 5 ANALYSIS:**

* We have a total of 78/110 audits with auditor details and these 78 audits have been done using 18 auditors.
* The challenge for CSQA audits is that the organization details are blank for 56/78 audits. Hence only performance metrics can be given for CSQA audits and a generalized manpower structure can be given.
* We observe that on an average it takes 135 days to complete a CSQA audit. This has been derived from the 23 completed audit cases.
* One important find is that among these 23 cases, 6 audits were completed by auditor CL with an average of 110 days.
* Hence with this insight we can say that in our distribution graph we have 1 audit carried out by single auditor which can be reduced to 4~5 audits per person to reduce manpower and utilize in the pending/in progress audits.

**ANALYTICS STAGE 6 : MODEL PROPOSAL**

The correlation plot is as below:



* From the above correlation plot we see a very low correlation of audit completion delta with other variables. Hence a regression model proposed with all these variables would be slow due to many dependent variables and less accurate due to low correlation.
* However, a strong correlation is found between Audit status with Audit Method, Proposed Quarter and Audit organization. Hence a classification model can be made to propose the audit status using decision trees or logistic regression however with only 40% of data availability , the model accuracy is questionable.

These are some of the major analytics and measures from the given dataset. However, for a more regulated audit qualification method it is advised to group nearby locations and assign single auditors for each groups and finally group data obtained from each auditor, to make sure there is no missing data and there is a regularized audit method. This is costly in the beginning but is an optimal solution in the long run.

**REFERENCE:**

**COMMITMENT TO RESEARCH AND DEVELOPMENT,** Retrieved from <https://www.vrtx.com/story/our-commitment-research-development>

**STOCK QUOTE VRTX 2019,** Retrieved from <https://investors.vrtx.com/stock-information>