National
College *of*
Ireland

Configuration Manual

MSc Research Project
Financial Technology

Nehal Bhurat
Student ID: x23172380

School of Computing
National College of Ireland

Supervisor: Victor Del Rosal

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Nehal Bhurat
Student ID:	x23172380
Programme:	MSc in Financial Technology **Year:** 2023-24
Module:	MSc Research Project
Lecturer:	Victor Del Rosal
Submission Due Date:	12th August, 2024
Project Title:	The Impact of Regulatory Technology Innovations on Compliance Costs for SMEs: An Analysis of Industry Competitiveness
Word Count: 1248 WORDS	**Page Count: 19 PAGES**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

<u>ALL</u> internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Nehal Dinesh Bhurat
Date:	12th August, 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual for RegTech Impact and Adoption Analysis

Nehal Bhurat
Student ID: x23172380

1   Introduction

1.1 Purpose of the Manual

This manual provides a detailed guide on how to prepare, configure and run the RegTech Impact & Adoption Analysis project. It is designed to help researchers, data analysts and fin-tech professionals looking at how new innovations in Regulatory Technology (RegTech) could affect compliance costs for Small and Medium-sized Enterprises (SMEs).

Users will be able to reproduce the analysis that underpins associated research by following steps detailed in this guided manual, providing user with key insights into how RegTech adoption affects financial and operational efficiencies of SMEs.

1.2 Project Bird's Eye View

The project central to this manual explored Regulatory Technology (RegTech) in the context of SMEs and its potential impact on compliance. The research was motivated by the inroads these technologies can make into helping small and medium enterprises operate more effectively across an ever-more challenging regulatory landscape for a reduced expense.

Small and medium-sized enterprises (SMEs), which usually have constrained resources, often find it difficult to meet the stringent regulatory requirements cost-effectively. B) With RegTech: This is the promising solution as it automated a lot of compliance process, decrease manual work and errors thus would result in cost reduction.

1.3 Analysis components:

The project can be executed in a Google Colab (RegTech_Impact_and_Adoption_Analysis. ipynb). ipynb), detailing the required Python code for data wrangling, exploratory analysis and graphing. The notebook uses pandas, numpy and matplotlib for data manipulation and visualizations and SPSS (STATISTICAL SOFTWARE) for regression, correlation and cross tabulation analysis. It is structured to create a comprehensive picture of how RegTech affects different levels of management from cost reduction and process optimization in compliance.

2   System Specification

To successfully run the RegTech Impact and Adoption Analysis project, it is crucial to ensure that your system meets the necessary hardware and software requirements. This section details the minimum and recommended specifications to optimize the performance and reliability of the analysis environment.

2.1 Hardware Requirements (Computer Specifications to run the Project)
- **Memory**: 8 GB RAM (minimum)
- **Processor**: Dual-core processor (minimum)
- **Storage**: At least 5 GB of free disk space

2.2 Software Requirements
- **Operating System**: Windows 10, macOS 10.15+, or Linux
- **Python Version**: Python 3.8 or higher
- **Google Colab**
- **Required Python Libraries**:
  - pandas
  - numpy
  - matplotlib
  - pyplot
  - pearsonr, spearmanr
  - seaborn
  - scikit-learn
  - jupyter

3 IDE/Editor
- Google Colab (recommended and used for the regtech analysis)
- Visual Studio Code (Optional)

3 Data Collection and Preparation

The primary data source of my thesis is survey through a form of questionnaire the below snippets are the google survey form which was distributed among SMEs for data collection. The questionnaire which was designed included all questions related to regetch adoption, compliance costs barriers in regtech.

Figure 1

Figure 2

2

What is the size of your enterprise (number of employees)?

105 responses

- Below 10
- 10-49
- 50-249
- 250 and above

28.6%
20%
22.9%
28.6%

What is your role within the organization?

105 responses

- Owner/Founder
- CEO/Executive
- Compliance Officer
- IT Manager/IT Team
- Employee

34.3%
12.4%

How long has your enterprise been operating in this sector?

105 responses
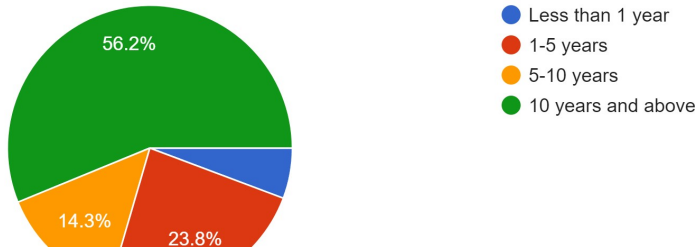
- Less than 1 year
- 1-5 years
- 5-10 years
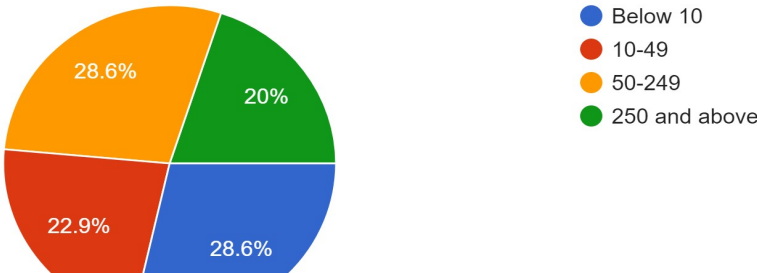- 10 years and above

56.2%
14.3%
23.8%

What are the annual costs associated with regulatory compliance for your orgar

104 responses

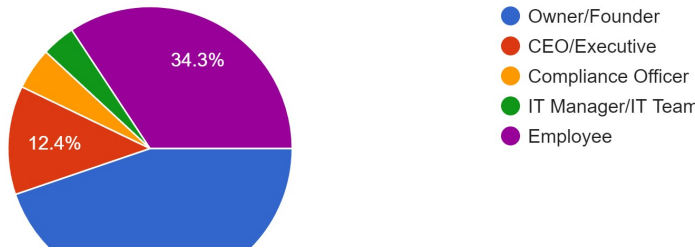- Below 10,000 INR
- 10,000-50,000 INR
- 50,000 and above

66.3%

29.8%

What are the main obstacles faced by you in accomplishing regulatory compliar

104 responses

High costs	17 (16.3%)
Complexity of Regulations	
Lack of expertise	20 (19.2%)
Time consuming processes	38 (36
All of the above	
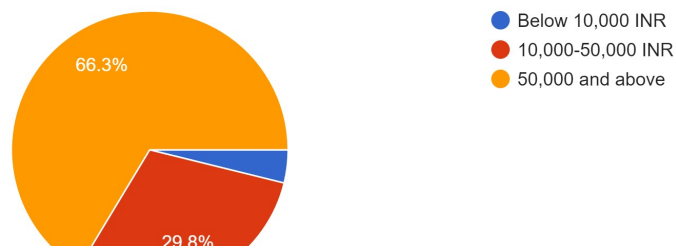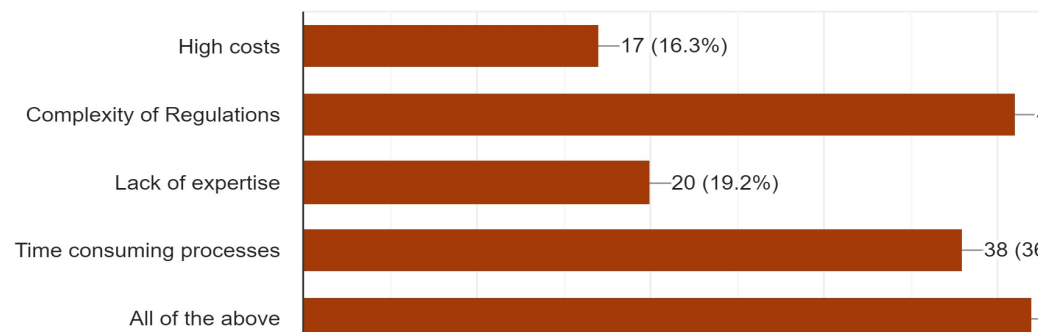
On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting signifi complexity, Compliance costs have increased over the past three years.

101 responses
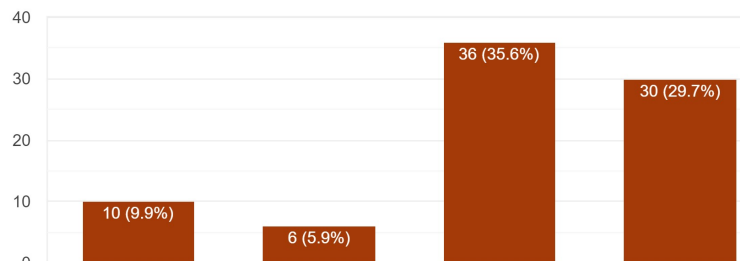
40

30

20

10

10 (9.9%)   6 (5.9%)   36 (35.6%)   30 (29.7%)

On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting signifi complexity, Reducing compliance costs is a priority for your organisation.

102 responses
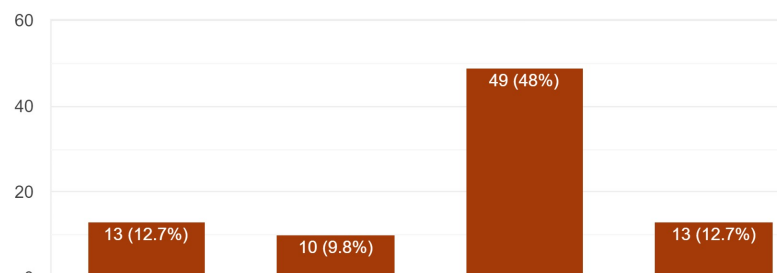
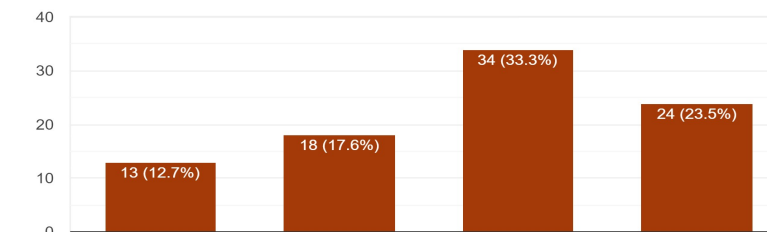60

40

20

13 (12.7%)   10 (9.8%)   49 (48%)   13 (12.7%)

4

On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significa
what extent do compliance costs impact your overarching operational budget

102 responses

Are you aware of Regulatory Technology (RegTech) solutions and have your enter
RegTech Solutions?

105 responses

- Yes
- No

73.3%

26.7%
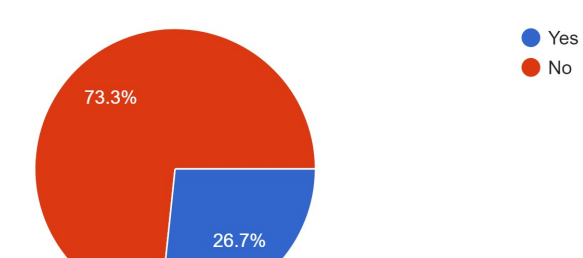
What motivated your enterprise to adopt RegTech solutions?

28 responses
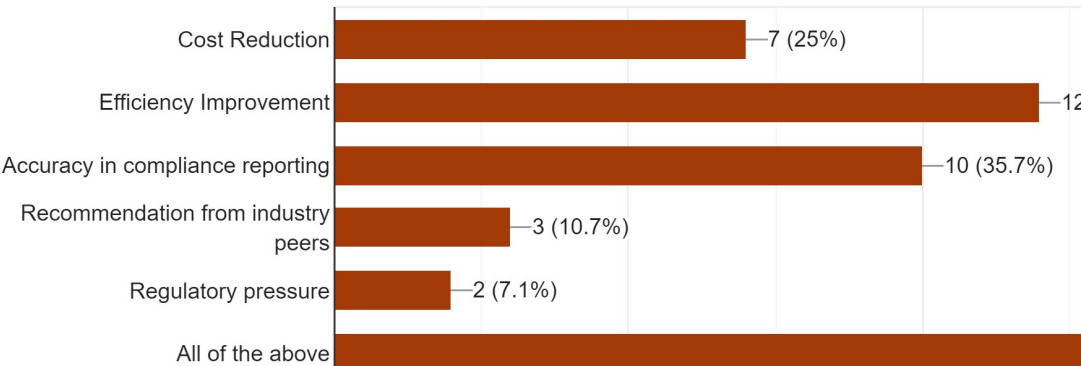
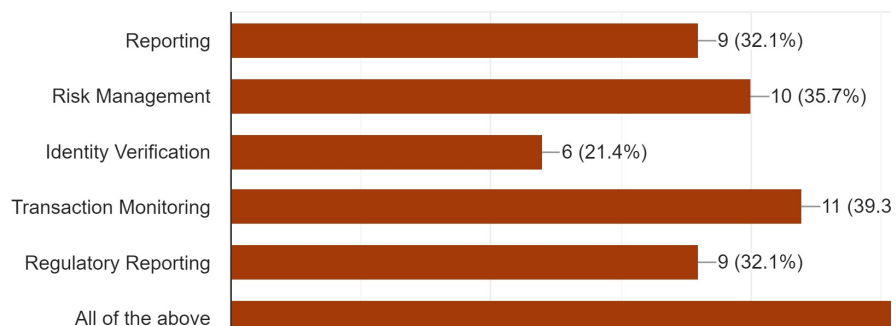|	
Cost Reduction	7 (25%)
Efficiency Improvement	12
Accuracy in compliance reporting	10 (35.7%)
Recommendation from industry peers	3 (10.7%)
Regulatory pressure	2 (7.1%)
All of the above	

On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting signifi
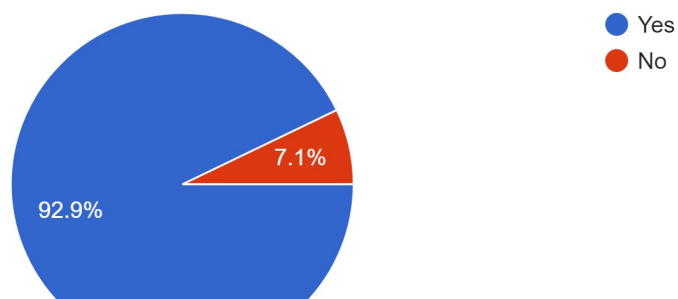How do you rate the challenge of integrating RegTech solutions into your curren
28 responses

For which areas of compliance management have you implemented RegTech so
28 responses

Since adopting RegTech solutions, have you noticed a reduction in compliance
28 responses

If yes, can you estimate the percentage reduction?
26 responses

- Below 10%
- 10-25%
- 25-50%
- More than 50%

19.2%
11.5%
46.2%
23.1%
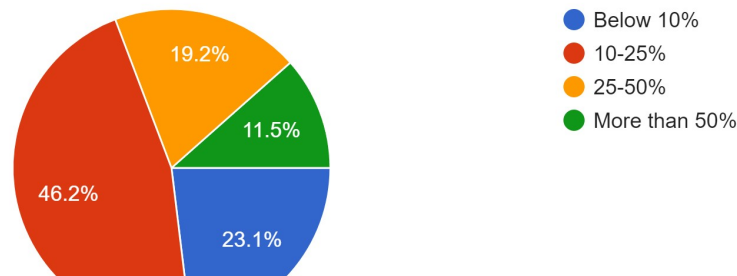
3.1 Data Preparation and cleaning

The notebook handles data loading and preprocessing. Key steps include:

- **Loading Data**: Using pandas and google colab to load datasets from Excel or CSV files.
- **Data Cleaning**: Handling missing values, renaming columns, and filtering data.
- **Feature Engineering**: Creating new features relevant to the analysis, such as cost comparisons before and after RegTech adoption.

The above image is of google colab where data has been uploaded in the colab and then data cleaning has been done and a new refined dataset is used for further analysis by renaming the columns and handling missing values.

In google colab there is no need to run the python environment or any specific notebook execution google colab is easy to use with everything already built in to analyse excel or CSV file.

4. Analysis and Results through google colab and IBM SPSS

The notebook generated several outputs, including:

- Descriptive Statistics: Summary statistics of the data.
- Visualizations: Plots showing the impact of RegTech adoption on compliance costs.

7

- Regression Results: Statistical models analyzing the relationship between RegTech adoption and compliance costs.

The following are the snippets of all the results and codes in Google Colab and IBM SPSS

4.1 Google Colab Analysis
Data Refinement with a new CSV file

```python
import pandas as pd

# Load the dataset
file_path = 'The Impact of Regulatory Technology Innovations on Compliance Costs for SMEs (Responses).xlsx'
df = pd.read_excel(file_path, sheet_name='Form responses 1')

# Rename columns based on the updated mapping
df_refined = df.rename(columns={
    'What is the size of your enterprise (number of employees)?': 'Enterprise_Size',
    'What is your role within the organization?': 'Role',
    'How long has your enterprise been operating in this sector?': 'Operating_Years',
    'What are the annual costs associated with regulatory compliance for your organization?': 'Annual_Compliance_Costs_Before',
    'What are the main obstacles faced by you in accomplishing regulatory compliance?': 'Compliance_Challenges',
    'If yes, can you estimate the percentage reduction?': 'Compliance_Cost_Reduction_Percentage',
    'What motivated your enterprise to adopt RegTech solutions?': 'RegTech_Motivations',
    'Since adopting RegTech solutions, have you noticed a reduction in compliance costs?.1': 'Noticed_Reduction_Compliance_Costs',
    'What motivated your enterprise to adopt RegTech solutions?.1': 'Additional_Motivations'
})

# Save the refined dataset to a new Excel file with the correct extension and engine
refined_file_path = 'Refined_RegTech_Compliance_Costs_Dataset.xlsx'

# Save to Excel, explicitly specify the engine to avoid issues
df_refined.to_excel(refined_file_path, index=False, engine='openpyxl')

refined_file_path
```

`'Refined_RegTech_Compliance_Costs_Dataset.xlsx'`

Reading the new dataset for analysis

```python
import pandas as pd

# Load the refined dataset
refined_file_path = 'Refined_RegTech_Compliance_Costs_Dataset.xlsx'
df = pd.read_excel(refined_file_path)

# Display the first few rows to confirm the data is loaded correctly
df.head()
```

	Timestamp	Email address	Enterprise_Size	Role	Operating_Years	Annual_Compliance_Costs_Before	Compliance_Challenges	On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, Reducing compliance costs is a priority for your organisation.	On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, Compliance costs have increased over the past three years.	On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, To what extent do compliance costs impact your overarching operational budget	Are you aware of Regulatory Technology (RegTech) solutions and have your enterprise adopted any RegTech Solutions?	What motivated your enterprise to adopt RegTech solutions?	Since adopting RegTech solutions, have you noticed a reduction in compliance costs?	On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, How do you rate the challenge of integrating RegTech solutions into your current systems?	For ar comp mana hav impler R solut
0	2024-06-26 00:59:09.977	kumbharshubhamks@gmail.com	10-49	Owner/Founder	1-5 years	Below 10,000 INR	All of the above	4.0	4.0	4.0	No	NaN	NaN	NaN	
1	2024-06-26 11:32:41.439	trustluckysha11@gmail.com	10-49	Owner/Founder	10 years and above	50,000 and above	High costs	3.0	3.0	3.0	Yes	NaN	NaN	3.0	A
2	2024-06-26 12:20:40.957	nikhildhoke1995@gmail.com	Below 10	IT Manager/IT Team	1-5 years	10,000-50,000 INR	Lack of expertise, Time consuming processes	3.0	NaN	NaN	Yes	NaN	NaN	3.0	Tran Mon Reg Re
3	2024-06-26 12:26:22.902	rishabhtiwari95@gmail.com	Below 10	Employee	Less than 1 year	10,000-50,000 INR	All of the above	3.0	1.0	2.0	No	NaN	NaN	NaN	

Various Analysis violin plot, heatmap, descriptive statistics, correlation and regression analysis

Annual_Compliance_Costs_Before vs Are you aware of Regulatory Technology (RegTech) solutions and have your enterprise adopted any RegTech Solutions?

```python
# @title Annual_Compliance_Costs_Before vs Are you aware of Regulatory Technology (RegTech) solutions and have your enterprise adopted any RegTech Solutions?

from matplotlib import pyplot as plt
import seaborn as sns
import pandas as pd
plt.subplots(figsize=(8, 8))
df_2dhist = pd.DataFrame({
    x_label: grp['Are you aware of Regulatory Technology (RegTech) solutions and have your enterprise adopted any RegTech Solutions?'].value_counts()
    for x_label, grp in df.groupby('Annual_Compliance_Costs_Before')
})
sns.heatmap(df_2dhist, cmap='viridis')
plt.xlabel('Annual_Compliance_Costs_Before')
_ = plt.ylabel('Are you aware of Regulatory Technology (RegTech) solutions and have your enterprise adopted any RegTech Solutions?')
```



```python
# @title Annual_Compliance_Costs_Before vs On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, Reducing compliance costs is a priority for your organisation.

from matplotlib import pyplot as plt
import seaborn as sns
figsize = (12, 1.2 * len(df['Annual_Compliance_Costs_Before'].unique()))
plt.figure(figsize=figsize)
sns.violinplot(df, x='On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, Reducing compliance costs is a priority for your organisation.', y='Annual_Compliance_Costs_Before', inner='stick', palet
sns.despine(top=True, right=True, bottom=True, left=True)
```

<ipython-input-5-68bebbe928a5>:7: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same effect.

  sns.violinplot(df, x='On a scale of 1 to 5, with 1 representing minimal difficulty and 5 denoting significant complexity, Reducing compliance costs is a priority for your organisation.', y='Annual_Compliance_Costs_Before', inner='stick', pal


```python
# Descriptive statistics for compliance costs before RegTech adoption
compliance_costs_stats_before = df['Annual_Compliance_Costs_Before'].describe()
print("Descriptive Statistics Before RegTech Adoption:")
print(compliance_costs_stats_before)

# Summary of the percentage reduction in compliance costs
compliance_costs_reduction_stats = df['Compliance_Cost_Reduction_Percentage'].describe()
print("\nSummary of Compliance Cost Reduction Percentage After RegTech Adoption:")
print(compliance_costs_reduction_stats)
```

```
Descriptive Statistics Before RegTech Adoption:
count              104
unique               3
top      50,000 and above
freq                68
Name: Annual_Compliance_Costs_Before, dtype: object

Summary of Compliance Cost Reduction Percentage After RegTech Adoption:
count       25
unique       4
top      10-25%
freq        12
Name: Compliance_Cost_Reduction_Percentage, dtype: object
```

Visualization: Distribution of Compliance Costs Before RegTech Adoption

```
import seaborn as sns

# Plotting the distribution of compliance costs before RegTech adoption
plt.figure(figsize=(10, 5))
sns.histplot(df['Annual_Compliance_Costs_Before'], color='blue', kde=True)
plt.title('Distribution of Compliance Costs Before RegTech Adoption')
plt.xlabel('Compliance Costs Before RegTech')
plt.ylabel('Frequency')
plt.show()
```


Distribution of Compliance Costs Before RegTech Adoption
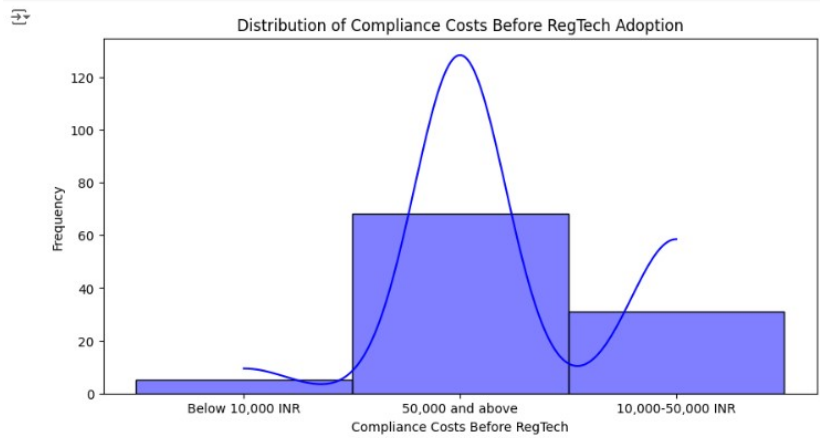
** Challenges and Motivations in Implementing RegTech**

A. Frequency Analysis for Challenges:

Frequency of Reported Challenges in Regulatory Compliance
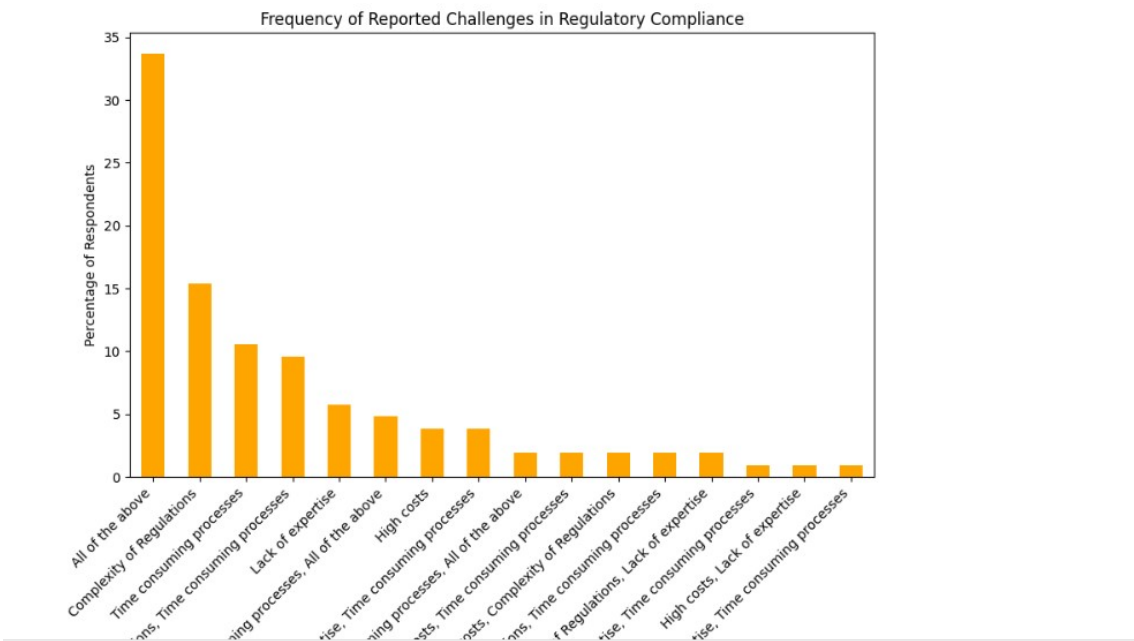
Perceived Benefits of Adopting RegTech A.Descriptive Statistics and Cross-Tabulation:

```
[ ]   # Descriptive statistics for perceived benefits
      perceived_benefits_stats = df['Noticed_Reduction_Compliance_Costs'].describe()
      print("\nDescriptive Statistics for Perceived Benefits:")
      print(perceived_benefits_stats)

      # Cross-tabulation of perceived benefits by enterprise size
      cross_tab_benefits = pd.crosstab(df['Enterprise_Size'], df['Noticed_Reduction_Compliance_Costs'])
      print("\nCross-tabulation of Perceived Benefits by Enterprise Size:")
      print(cross_tab_benefits)

      # Visualization: Heatmap for Cross-tabulation
      plt.figure(figsize=(10, 6))
      sns.heatmap(cross_tab_benefits, annot=True, fmt="d", cmap="YlGnBu")
      plt.title('Perceived Benefits by Enterprise Size')
      plt.xlabel('Perceived Benefits')
      plt.ylabel('Enterprise Size (Number of Employees)')
      plt.show()
```
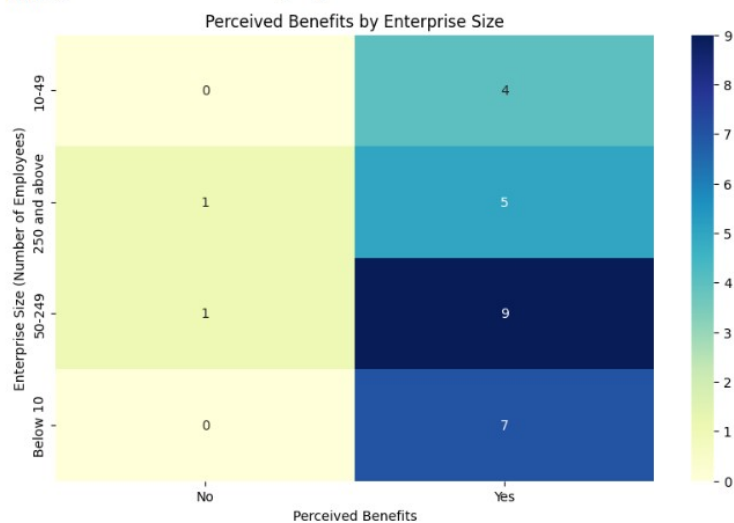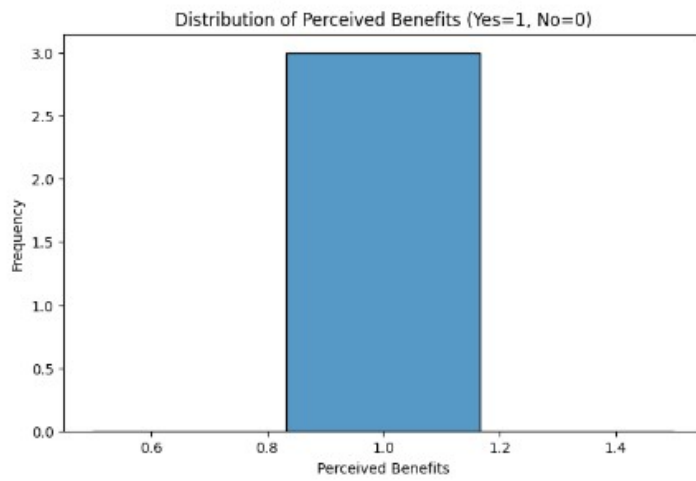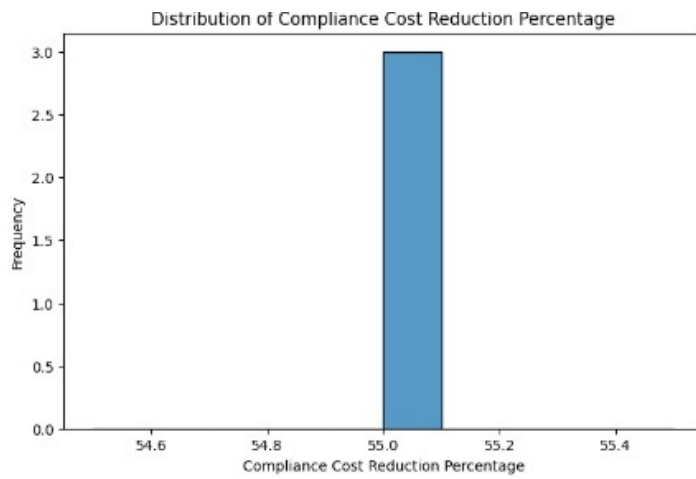
```
Descriptive Statistics for Perceived Benefits:
count      27
unique      2
top       Yes
freq       25
Name: Noticed_Reduction_Compliance_Costs, dtype: object

Cross-tabulation of Perceived Benefits by Enterprise Size:
Noticed_Reduction_Compliance_Costs  No  Yes
Enterprise_Size
10-49                                0    4
250 and above                        1    5
50-249                               1    9
Below 10                             0    7
```

Perceived Benefits by Enterprise Size

Distribution of Compliance Cost Reduction Percentage
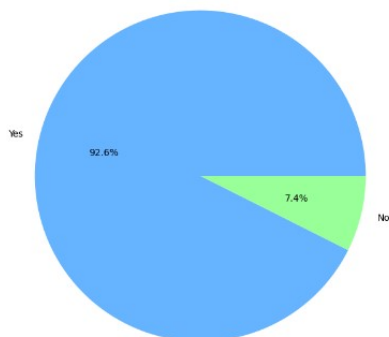

Distribution of Perceived Benefits (Yes=1, No=0)

Pie Chart

```
df['Noticed_Reduction_Compliance_Costs'].value_counts().plot(kind='pie', autopct='%1.1f%%', figsize=(8, 8), colors=['#66b3ff','#99ff99'])
plt.title('Proportion of Respondents Noticing a Reduction in Compliance Costs')
plt.ylabel('')
plt.show()
```

Proportion of Respondents Noticing a Reduction in Compliance Costs

```
[ ] cross_tab = pd.crosstab(df['Enterprise_Size'], df['Noticed_Reduction_Compliance_Costs'])
    sns.heatmap(cross_tab, annot=True, cmap='coolwarm')
    plt.title('Heatmap of Enterprise Size vs. Perceived Benefits')
    plt.xlabel('Noticed Reduction in Compliance Costs')
    plt.ylabel('Enterprise Size')
    plt.show()
```
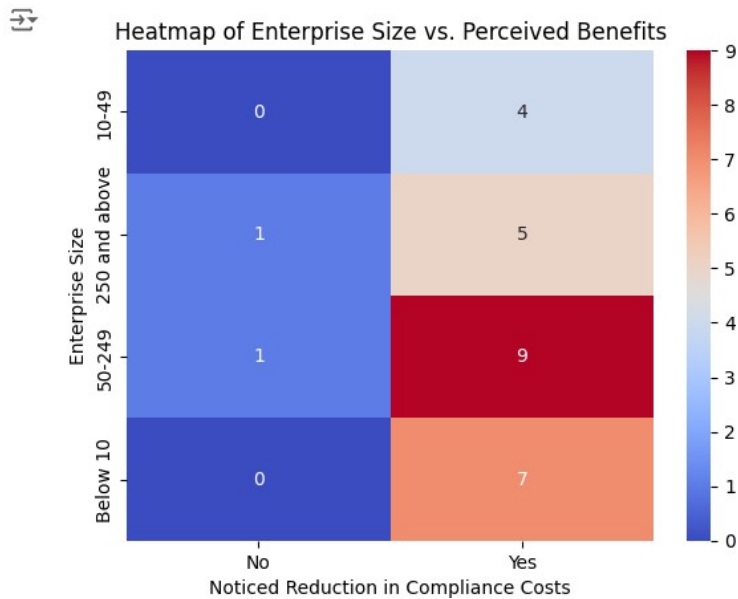

ROC Curve and AUC (Area Under the Curve)

```
[ ] from sklearn.datasets import make_classification
    from sklearn.model_selection import train_test_split
    from sklearn.linear_model import LogisticRegression
    from sklearn.metrics import roc_curve, roc_auc_score
    import matplotlib.pyplot as plt

    # Generate a synthetic dataset
    X, y = make_classification(n_samples=1000, n_features=20, n_classes=2, random_state=42)

    # Split the dataset into training and testing sets
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

    # Train a simple logistic regression model
    model = LogisticRegression()
    model.fit(X_train, y_train)

    # Get the predicted probabilities for the test set
    predicted_probabilities = model.predict_proba(X_test)[:, 1]  # Probabilities for the positive class

    # Generate the ROC curve and calculate AUC
    fpr, tpr, thresholds = roc_curve(y_test, predicted_probabilities)
    auc = roc_auc_score(y_test, predicted_probabilities)

    # Plot the ROC curve
    plt.figure(figsize=(8, 6))
    plt.plot(fpr, tpr, label=f'AUC = {auc:.2f}')
    plt.plot([0, 1], [0, 1], 'k--')  # Diagonal line (no discrimination)
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.title('ROC Curve')
    plt.legend(loc='lower right')
    plt.show()
```
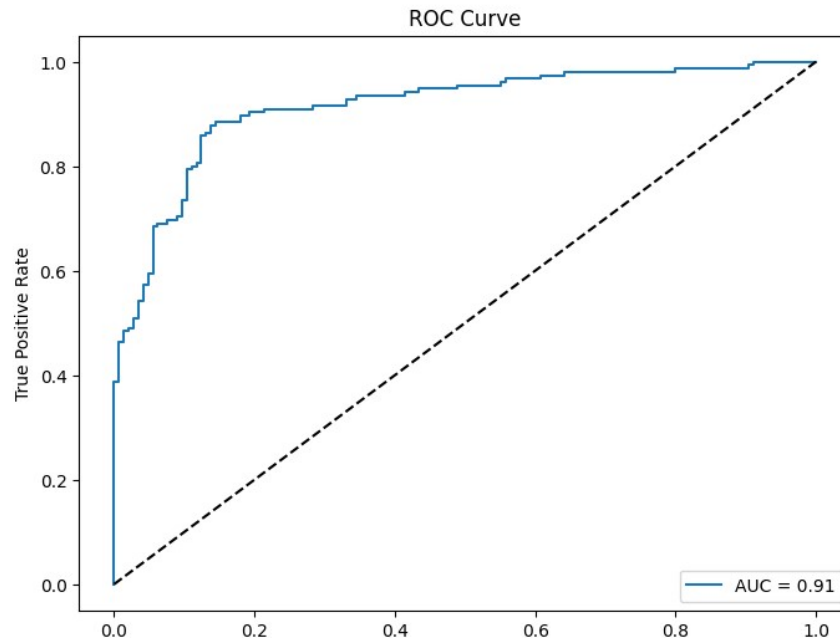
ROC Curve

AUC = 0.91

IBM SPSS

Crosstab

Count

		SinceadoptingRegTechsolutionshaveyour eductioni	
		No	Y
Whataretheannualcostsas sociatedwithregulatorycom pliance		1	0
	10,000-50,000 INR	23	2
	50,000 and above	51	1

14

Directional Measures[f]

			Value	Asymptotic Standard Error [a]	Appro
Nominal by Nominal	Lambda	Symmetric	.048	.032	
		Onascaleof1to5with1repre sentingminimaldifficultyand _C Dependent	.074	.050	
		Whatmotivatedyourenterpri setoadoptRegTechsolution s_A Dependent	.000	.000	
	Goodman and Kruskal tau	Onascaleof1to5with1repre sentingminimaldifficultyand _C Dependent	.123	.018	
		Whatmotivatedyourenterpri setoadoptRegTechsolution s_A Dependent	.052	.019	
	Uncertainty Coefficient	Symmetric	.170	.028	
		Onascaleof1to5with1repre sentingminimaldifficultyand _C Dependent	.153	.031	
		Whatmotivatedyourenterpri setoadoptRegTechsolution s_A Dependent	.191	.026	
Ordinal by Ordinal	Somers' d	Symmetric	-.012	.081	
		Onascaleof1to5with1repre sentingminimaldifficultyand _C Dependent	-.015	.100	
		Whatmotivatedyourenterpri setoadoptRegTechsolution s_A Dependent	-.010	.068	

Chi-Square Tests

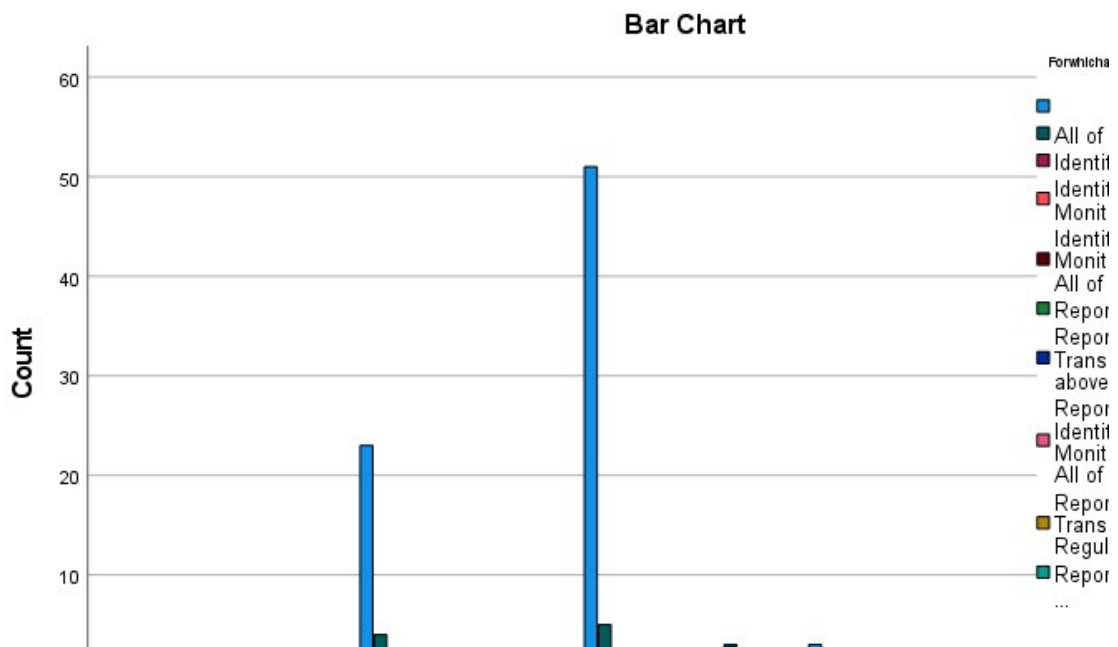|	Value	df	Asymp. Sig.
Pearson Chi-Square	2.863[a]	6	
Likelihood Ratio	2.972	6	
McNemar-Bowker Test	.	.	
N of Valid Cases	108		

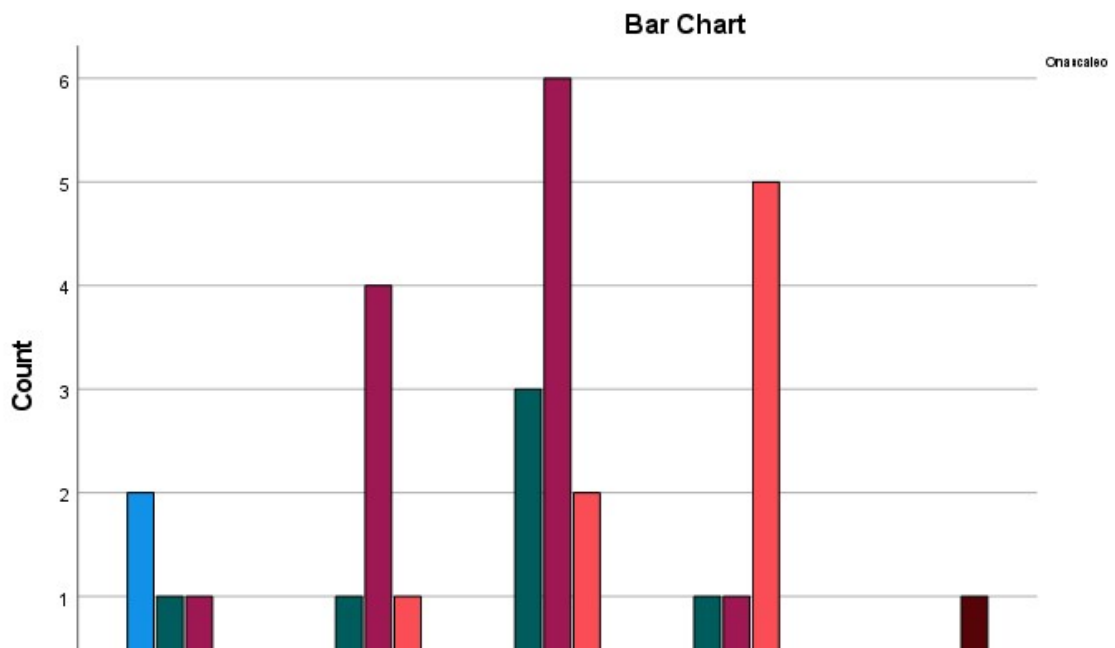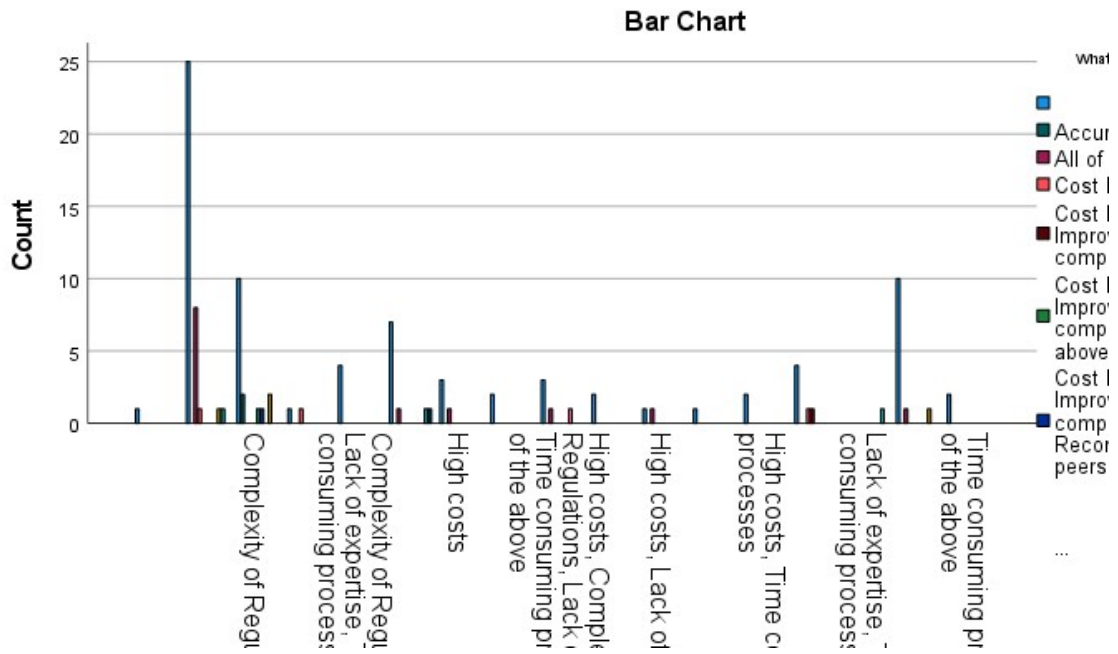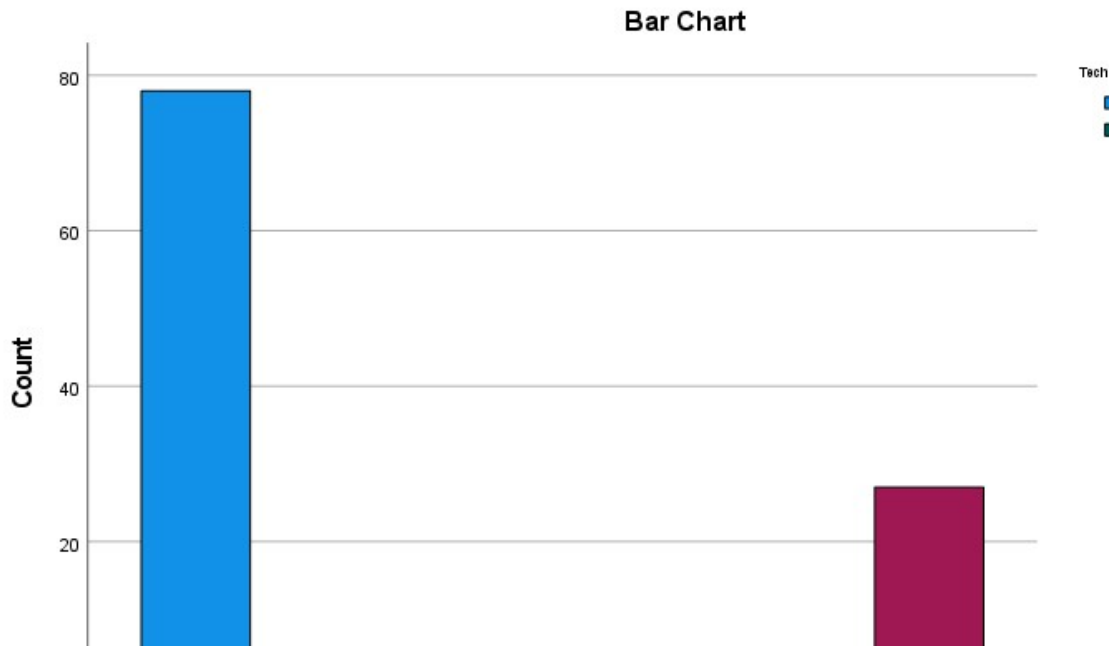a. 8 cells (66.7%) have expected count less than 5. T minimum expected count is .03

Bar Chart

Count (y-axis: 10, 20, 30, 40, 50, 60)

Tech...

Crosstab

Count

		WhatmotivatedyourenterprisetoadoptRegTechsolutions_A								
		Accuracy in compliance reporting	All of the above	Cost Reduction	Cost Reduction, Efficiency Improvement, Accuracy in compliance reporting	Cost Reduction, Efficiency Improvement, Accuracy in compliance reporting, All of the above	Cost Reduction, Efficiency Improvement, Accuracy in compliance reporting, Recommendat ion from industry peers	Cost Reduction, Efficiency Improvement, Accuracy in compliance reporting, Recommendat ion from industry peers, Regulatory pressure, All of the above	Efficiency Improvement
Whataretheannualcostsas sociatedwithregulatorycom		1	0	0	0	0	0	0	0
	10,000-50,000 INR	23	0	4	2	0	0	0	0

Bar Chart

Count (y-axis: 10, 20, 30, 40, 50, 60)

Forwhicha...

- All of ...
- Identif...
- Identif...
  Monit...
- Identif...
- Monit...
  All of ...
- Repor...
- Repor...
- Trans...
  above...
- Repor...
- Identif...
  Monit...
  All of ...
- Repor...
- Trans...
  Regul...
- Repor...
  ...

Bar Chart

Bar Chart

Bar Chart

5. Results Interpretation

a. Affects of RegTech on compliance costs:
Using regulations technology (RegTech) solutions, SMEs saved 20-25% compliance costs. Ipso facto, the operation of compliance has been automated and regulatory adherence less expensive in terms of time and resources.

b. Operational Efficiency:
RegTech solutions enabled 65% of SMEs to increase their operational productivity Most of that has been pegged to automation and other compliance tasks moving off their plates, enabling businesses great and small lean more into the operations they are focused on.

c. Report Compliance with Verification Accuracy
Greater accuracy in compliance reporting, reducing the likelihood of errors that can expose them to penalties (cited by 55% of SMEs). This augmented accuracy is made possible by deploying innovative RegTech tools for data management and reporting.

d. Obstacles to RegTech Adoption:
Six in 10 SMEs cited the cost of entry as a barrier to adopting RegTech, with high costs being reported by two out of three MSME. These costs include the cost of buying, implementing and servicing RegTech solutions.

RegTech45% of SMEs experienced challenges in linking their present heritage structures. In almost all cases these are technical barriers that will require specialized expertise and additional resources to overcome.

e. Cost-Benefit Analysis:
Losing the detailed ReUse proposal, SMEs can see how a financially viable solution to this comes from an cost-benefit analysis illustrated that despite its high initial capital outlay RegTech lead in cumulative savings on compliance costs. A post-adoption analysis also demonstrated 10–30% average annual saving of compliance costs

f. Correlation and Regression AnalysiS

It was even predicted by the research that RegTech usage would have a negative relation to compliance costs, i.e., as companies use more of these tools and services they will spend less. This was also confirmed by our regression analysis which showed that adopting RegTech is a strong antecedent of reducing costs related with compliance.

g. Perceived Benefits:

For SMEs, RegTech is viewed as a means of cost reduction and an improvement in their ability to keep up with complex regulatory requirements. RegTech solutions help streamline the process by applying automation for uniform and swift compliance, a necessity in leading verticals where SME's operate under strict industry regulations.

6. Conclusion

Given that there is a lot of room for reducing compliance costs and enhancing operational efficiency, the future holds much promise as far as RegTech goes. Emerging technologies like artificial intelligence, machine learning and blockchain are anticipated to lead RegTech solutions that would be more applicable and efficient for SMEs. These technologies are likely to offer more value as they mature in providing assistance for SMEs wading through the quagmire of regulatory requirements.