

Configuration Manual

1. Introduction and Purpose

This Configuration Manual provides detailed instructions on setting up and configuring the software, tools, and systems used in the thesis titled "Multi-Cloud: Assessing Resilience Amid Threats." The manual is intended for researchers, IT professionals, and students who wish to replicate the study, verify the results, or apply the developed models and methods to their own work.

2. System Requirements

2.1 Hardware Requirements

- **Processor:** Intel i5 or equivalent
- **RAM:** 8 GB minimum, used – 16 GB used
- **Storage:** 20 GB of free disk space
- **Network:** Broadband internet connection for downloading software and accessing cloud services

2.2 Software Requirements

- **Operating System:** Windows 10, macOS, or Linux – Windows 10 used
- **IBM SPSS Statistics:** Version 26 or higher
- **Python:** Version 3.7 or higher
- **Visual Studio Code:** For Python code review and execution
- **Python Libraries:** pandas, numpy, matplotlib.pyplot, seaborn

2.3 Network Requirements

- **Internet Access:** Required for software downloads, updates, and accessing online resources
- **Firewall Settings:** Ensure necessary ports for internet access are open

3. Installation Instructions

3.1 Installing IBM SPSS Statistics

1. **Download SPSS:** Visit the IBM website (<https://www.ibm.com/products/spss-statistics>) and download the appropriate version of IBM SPSS Statistics for your operating system.
2. **Install SPSS:** Follow the installation prompts to install SPSS on your machine.
3. **License Activation:** Activate the software using the license key provided by your institution or purchase a license through IBM.

3.2 Installing Python and Visual Studio Code

1. Download Python: Visit the Python website (<https://www.python.org/downloads/>) and download the latest version of Python.
2. Install Python: Follow the installation prompts, ensuring that you check the option to "Add Python to PATH".
3. Download Visual Studio Code: Visit the Visual Studio Code website (<https://code.visualstudio.com/>) and download the latest version for your operating system.
4. Install Visual Studio Code: Follow the installation prompts to install Visual Studio Code.
5. Install Python Extension for Visual Studio Code:
 - Open Visual Studio Code.
 - Go to the Extensions view by clicking on the Extensions icon in the Activity Bar on the side of the window.
 - Search for "Python" and install the Python extension by Microsoft.

3.3 Installing Required Python Libraries

Run the following commands in your terminal or command prompt:

```
pip install pandas numpy matplotlib seaborn
```

4. System Configuration

4.1 Initial Configuration

1. **SPSS Configuration:**
 - Open IBM SPSS Statistics.
 - Set your working directory to the folder containing your dataset and SPSS syntax files.
 - Load your dataset into SPSS by navigating to "File > Open > Data" and selecting the appropriate file.
 - Switch to the variable view to configure the variable as scale, nominal or ordinal. Also set your decimals appropriately
2. **Python Environment Configuration:**
 - Ensure that all necessary Python libraries are installed by running the provided installation commands.

4.2 Customization

- **SPSS Syntax:** Modify the provided SPSS syntax files to reflect your specific dataset locations and parameters.
- **Python Scripts:** Adjust the Python scripts to match your data processing and analysis needs.

4.3 Testing the Setup

1. **Run Test Scripts:**
 - Use the provided SPSS syntax files to run test analyses and verify that the environment is correctly configured.
2. **Check Output:**
 - Verify that the output from test scripts and SPSS matches expected results, such as correctly calculated resilience scores and visualizations.

5. Usage Instructions

5.1 Running the Analysis

1. **IBM SPSS Analysis & Statistics:**
 - Load your dataset into SPSS.
 - Open the provided SPSS syntax file (Enterprise_Resilience.sav).
 - On SPSS main menu, go to Analyze > Regression > Linear
 - In the pop up, select the dependent variable(Resilience Score) and populate the independent variables(predictors) in the predictors block
 - Click Statistics and make sure “Estimates, Model Fit and R Squared change” are selected and click “Continue”.
 - Click Plot and from the DEPENDENT block, drag “*ZPRED” to the X-axis and the “*ZRESID” to the Y-axis. Click “Continue” and click “Ok”
 - This will run and perform the multiple linear regression test.
 - View and export the results as needed.

5.2 Advanced Features

- **Custom Data Inputs:** Modify the CSV files or SPSS datasets with your data to run custom analyses.
- **Parameter Tuning:** Adjust regression model parameters within the SPSS syntax or Python scripts to test different scenarios.

5.3 Troubleshooting

- Make sure your variables are well defined in the variable view
- **Common Errors:** Refer to the troubleshooting section for solutions to common issues such as SPSS syntax errors, Python library installation problems, or script execution failures.
- **Error Logs:** Review error logs generated by SPSS and Visual Studio Code for detailed information.

6. Maintenance and Updates

6.1 Regular Maintenance

- **SPSS and Python Updates:** Regularly update IBM SPSS and Python libraries to the latest versions to ensure compatibility and access to new features.

- **File Management:** Organize and back up your SPSS datasets, syntax files, and Python scripts regularly to prevent data loss..

Appendix

```

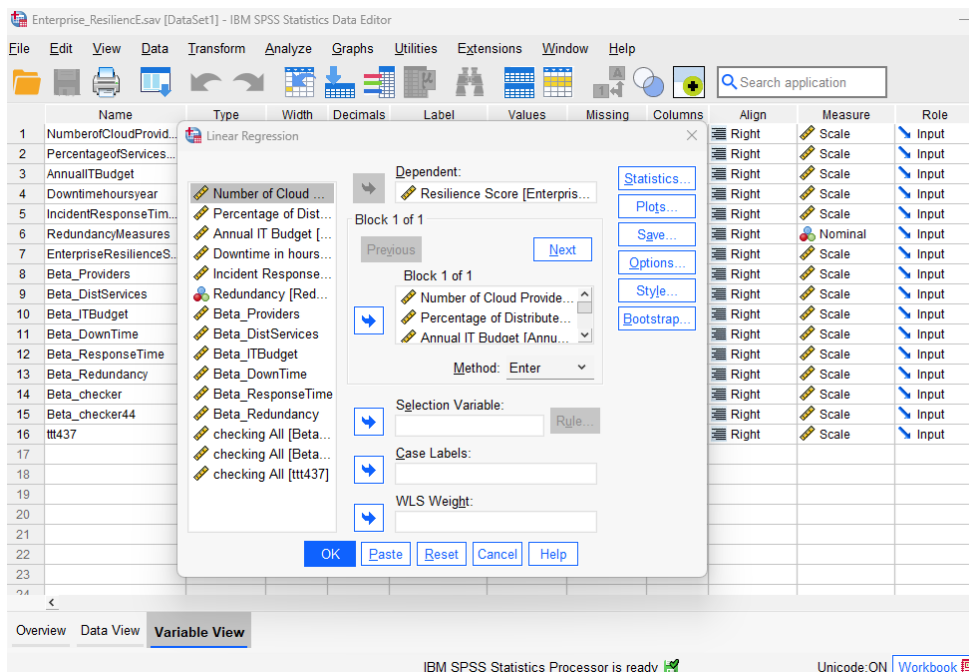
1  import pandas as pd
2  import numpy as np
3
4  # Load the existing data
5  file_path = 'C:\Users\LineX\Downloads\enterprise_resilience.csv'
6  data = pd.read_csv(file_path)
7
8  # Define the number of new rows needed
9  num_new_rows = 1000 - len(data)
10
11 # Generate synthetic data
12 np.random.seed(42) # For reproducibility
13 new_data = pd.DataFrame({
14     'Number of Cloud Providers': np.random.choice(data['Number of Cloud Providers'], num_new_rows),
15     'Percentage of Services Distributed': np.random.choice(data['Percentage of Services Distributed'], num_new_rows),
16     'Annual IT Budget': np.random.choice(data['Annual IT Budget'], num_new_rows),
17     'Downtime (hours/year)': np.random.choice(data['Downtime (hours/year)'], num_new_rows),
18     'Incident Response Time (hours)': np.random.choice(data['Incident Response Time (hours)'], num_new_rows),
19     'Redundancy Measures': np.random.choice(data['Redundancy Measures'], num_new_rows),
20     'Enterprise Resilience Score': np.random.choice(data['Enterprise Resilience Score'], num_new_rows)
21 })
22
23 # Combine the original and synthetic data
24 augmented_data = pd.concat([data, new_data], ignore_index=True)
25
26 # Save the augmented dataset
27 augmented_data.to_csv('C:\Users\LineX\Downloads\aug_enterprise_resilience.csv', index=False)
28
29 # Display the first few rows of the augmented data to verify
30 print(augmented_data.head())
31 print("New dataset size:", len(augmented_data))
32

```

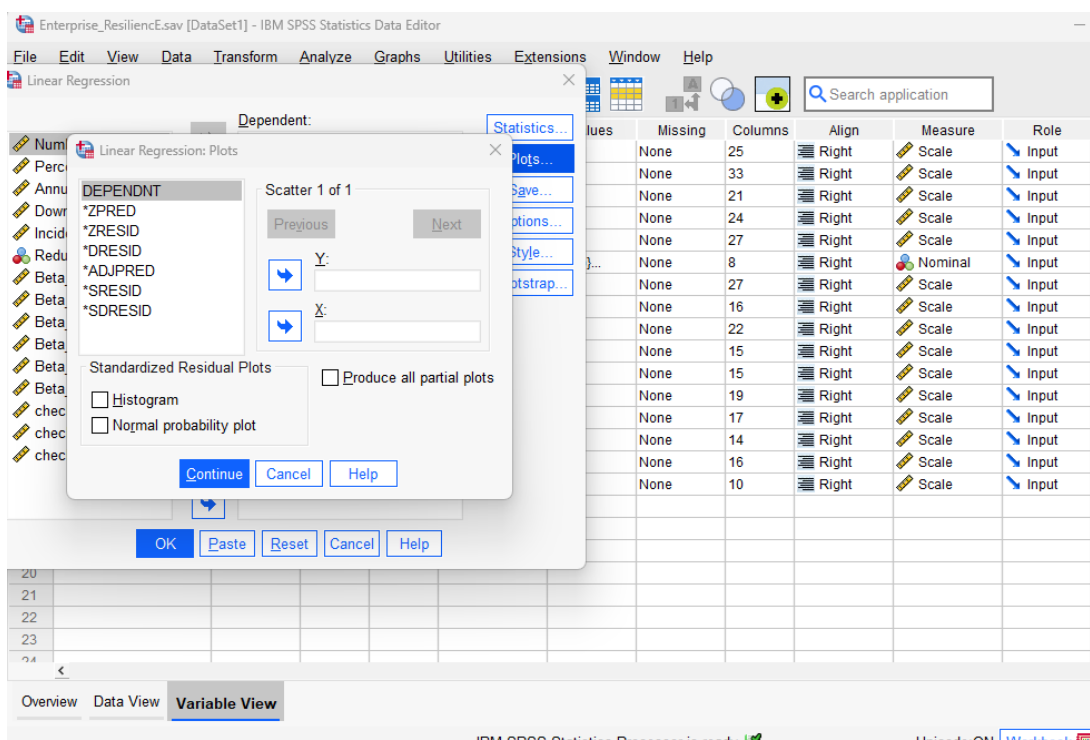
Python Script for data augmentation

	Number of Cloud Providers	Percentage of Services Distributed	Annual IT Budget	Downtime hours/year	Incident Response Time hours	Redundancy Measure
1	3	75.8529	1086255.3918	82.7519	7.5899	
2	4	62.8877	1010445.5675	76.4528	19.4607	
3	1	44.7622	148719.9250	57.3529	23.9365	
4	3	85.1036	748370.8725	95.6047	1.6906	
5	3	74.7785	822371.6757	20.0475	21.6394	
6	4	33.0094	857763.2837	10.9284	15.3205	
7	1	92.8742	1202327.5015	85.3962	23.3878	
8	1	85.8030	1113844.8388	43.9150	11.6952	
9	3	95.9840	1255019.6763	84.6969	20.4899	
10	2	78.0576	1553278.1969	89.3090	2.4351	
11	3	69.0732	1644672.9036	6.2458	8.7114	
12	3	53.4594	1464433.8448	88.3465	2.5414	
13	3	94.6183	1915495.0223	44.8319	23.4329	
14	3	89.2851	134641.9071	51.0432	19.7904	
15	4	23.6175	471978.1728	62.6593	20.6086	
16	1	22.1094	114369.4625	92.6383	22.5732	
17	4	50.1171	1330201.9569	1.9116	2.9574	
18	4	84.8443	1806258.0430	47.6843	9.8687	
19	4	98.9821	562616.3635	68.7722	2.6338	
20	3	32.0334	1861365.6351	72.2707	5.8482	
21	2	67.5105	214508.0416	69.2512	6.2754	

SPSS Data view



Configuring dependent variable and independent variables in SPSS



SPSS Plot setting

SPSS Output panel

IBM SPSS Statistics Processor is ready Unicode: ON

Workbook2 [Document2] - IBM SPSS Statistics Workbook

File Edit View Data Transform Insert Run Format Analyze Graphs Utilities Extensions Window Help

Active DataSet: DataSet1

Budget

a. Dependent Variable: Resilience Score
b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.982 ^a	.964	.963	4.9985456	.964	4372.017	6	993	< .001

a. Predictors: (Constant), Redundancy, Number of Cloud Providers, Downtime in hours/year, Percentage of Distributed Services, Incident Response Time in hours, Annual IT Budget

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	655421.143	6	109236.857	4372.017	< .001 ^b
	Residual	24810.560	993	24.985		
	Total	680231.703	999			

a. Dependent Variable: Resilience Score
b. Predictors: (Constant), Redundancy, Number of Cloud Providers, Downtime in hours/year, Percentage of Distributed Services, Incident Response Time in hours, Annual IT Budget

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.
		B	Std. Error	Beta	1		
1	(Constant)	.369	.754		.499	.626	
	Number of Cloud Providers	9.891	.137	.437	71.876	< .001	
	Percentage of Distributed Services	.169	.007	.169	27.891	< .001	
	Annual IT Budget	9.949E-6	.006	.207	34.133	< .001	
	Downtime in hours/year	.496	.006	-.550	90.716	< .001	
	Incident Response Time in hours	-3.006	.024	-.606	93.394	< .001	
	Redundancy	20.427	.317	.392	64.536	< .001	

a. Dependent Variable: Resilience Score