

CSRIC_Analysis_Report.Rmd

Elvira Khwatenge

2024-11-07

R Markdown

be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
# Load necessary libraries
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)

# Load the dataset (update the path as necessary)
csric_data <- read.csv("CSRIC_Best_Practices_Raw.csv.csv")

# Preview the data
head(csric_data)
```

```
##   BP.Number Priority
## 1 11-10-5064
## 2 11-10-5164
## 3 11-11-3249
## 4  11-7-8548
## 5  11-8-8051
## 6  11-8-8052
##
## 1
## 2
## 3
## 4
## 5
## 6 "Retired" - Network Operators should mitigate limited SS7 authentication by enabling logging for S
##           Network.Type.s.
## 1 Cable; Internet/Data; Satellite; Wireless; Wireline;
## 2 Cable; Internet/Data; Satellite; Wireless; Wireline;
## 3 Cable; Internet/Data; Satellite; Wireless; Wireline;
```

```

## 4 Cable; Internet/Data; Satellite; Wireless; Wireline;
## 5                               Wireline;
## 6                               Wireline;
##                               Industry.Role.s.
## 1   Service Provider; Network Operator; Public Safety; Property Manager;
## 2 Service Provider; Network Operator; Equipment Supplier; Public Safety;
## 3                               Public Safety;
## 4                               Service Provider; Network Operator;
## 5   Service Provider; Network Operator; Equipment Supplier;
## 6                               Network Operator;
##                               Keywords
## 1                               Buildings; Fire; Network Operations;
## 2   Corporate Ethics; Physical Security Management; Policy; Supervision;
## 3
## 4                               Cyber Security; Disaster Recovery;
## 5   Cyber Security; Encryption; Network Elements; Network Operations; Policy;
## 6 Cyber Security; Intrusion Detection; Network Elements; Network Operations;
##   Public.Safety.and.Disaster
## 1                               TRUE
## 2                               TRUE
## 3                               TRUE
## 4                               FALSE
## 5                               FALSE
## 6                               FALSE
##
## 1   Reference: GR63 NEBS Requirements: Physical Protection, Telcordia, http://telecom-info.telcordia.com;
## 2
## 3
## 4                               IETF RFC2350, US-CERT\NRIC BP 8074, 8075, 0561, 0562;
## 5   ITU SS7 Standards, "Securing SS7 Telecommunications Networks", Proceedings of the 2001 IEEE Workshop on
## 6   Security and Privacy in the Presence of Untrusted Partners in the Internet;
##   cable internet.Data satellite wireless wireline Service.Provider
## 1   TRUE         TRUE         TRUE         TRUE         TRUE         TRUE
## 2   TRUE         TRUE         TRUE         TRUE         TRUE         TRUE
## 3   TRUE         TRUE         TRUE         TRUE         TRUE         FALSE
## 4   TRUE         TRUE         TRUE         TRUE         TRUE         TRUE
## 5   FALSE        FALSE        FALSE        FALSE        TRUE         TRUE
## 6   FALSE        FALSE        FALSE        FALSE        TRUE         FALSE
##   Network.Operator Priority..1.2.3. Equipment.Supplier Property.Manager
## 1           TRUE                NA                FALSE        TRUE
## 2           TRUE                NA                TRUE         FALSE
## 3           FALSE               NA                FALSE        FALSE
## 4           TRUE                NA                FALSE        FALSE
## 5           TRUE                NA                TRUE         FALSE
## 6           TRUE                NA                FALSE        FALSE
##   Government Public.Safety
## 1           FALSE              TRUE
## 2           FALSE              TRUE
## 3           FALSE              TRUE
## 4           FALSE              FALSE
## 5           FALSE              FALSE
## 6           FALSE              FALSE

```

Data cleaning

3.1 Checking for Missing Values

```
##          BP.Number          Priority
##          0              0
##      Description      Network.Type.s.
##          0              0
##      Industry.Role.s.      Keywords
##          0              0
## Public.Safety.and.Disaster      Reference
##          0              0
##          cable      internet.Data
##          0              0
##      satellite      wireless
##          0              0
##      wireline      Service.Provider
##          0              0
##      Network.Operator      Priority..1.2.3.
##          0              123
##      Equipment.Supplier      Property.Manager
##          0              0
##          Government      Public.Safety
##          0              0
```

3.2 Removing Duplicates and Setting Data Types

```
# Remove duplicates
csric_data <- csric_data[!duplicated(csric_data), ]

# Convert relevant columns to factors for categorical analysis
csric_data$Priority <- as.factor(csric_data$Priority)
# Print column names to confirm exact names
colnames(csric_data)

## [1] "BP.Number"          "Priority"
## [3] "Description"        "Network.Type.s."
## [5] "Industry.Role.s."   "Keywords"
## [7] "Public.Safety.and.Disaster" "Reference"
## [9] "cable"              "internet.Data"
## [11] "satellite"          "wireless"
## [13] "wireline"           "Service.Provider"
## [15] "Network.Operator"   "Priority..1.2.3."
## [17] "Equipment.Supplier" "Property.Manager"
## [19] "Government"         "Public.Safety"

csric_data$Public.Safety.and.Disaster <-as.factor(csric_data$Public.Safety.and.Disaster)
csric_data$Network.Operator <- as.factor(csric_data$Network.Operator)
csric_data$Industry_Role.s. <- as.factor(csric_data$Industry.Role.s.)
```

4. Exploratory Data Analysis (EDA) 4.1 Summary Statistics

```
# Display summary statistics for key columns
summary(csric_data)

## BP.Number          Priority      Description
## Length:1129        :123      Length:1129
## Class :character    Critical    :191      Class :character
```

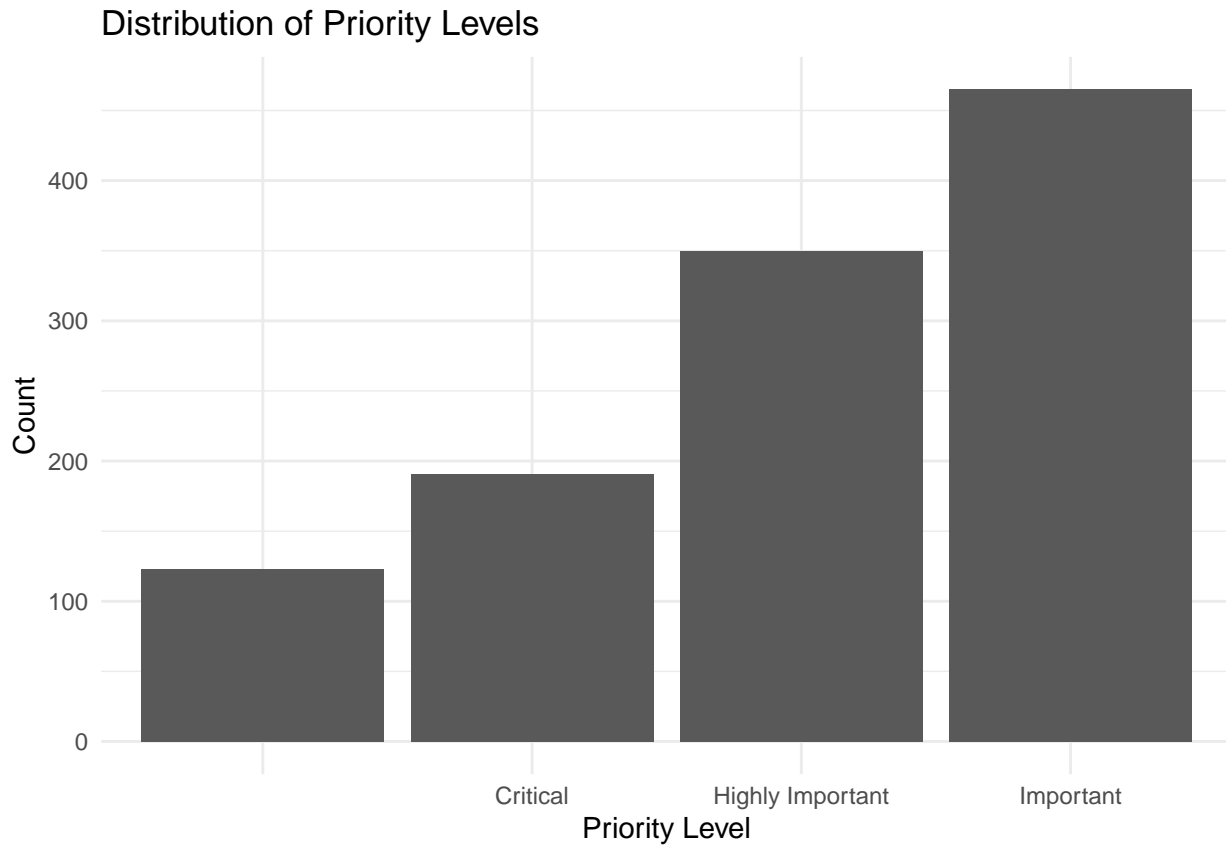
```

## Mode :character Highly Important:350 Mode :character
## Important :465
##
##
##
## Network.Type.s. Industry.Role.s. Keywords
## Length:1129 Length:1129 Length:1129
## Class :character Class :character Class :character
## Mode :character Mode :character Mode :character
##
##
##
## Public.Safety.and.Disaster Reference cable internet.Data
## FALSE:513 Length:1129 Mode :logical Mode :logical
## TRUE :616 Class :character FALSE:228 FALSE:109
## Mode :character TRUE :901 TRUE :1020
##
##
##
## satellite wireless wireline Service.Provider
## Mode :logical Mode :logical Mode :logical Mode :logical
## FALSE:353 FALSE:216 FALSE:232 FALSE:257
## TRUE :776 TRUE :913 TRUE :897 TRUE :872
##
##
##
## Network.Operator Priority..1.2.3. Equipment.Supplier Property.Manager
## FALSE:183 Min. :1.000 Mode :logical Mode :logical
## TRUE :946 1st Qu.:1.000 FALSE:725 FALSE:947
## Median :2.000 TRUE :404 TRUE :182
## Mean :1.728
## 3rd Qu.:2.000
## Max. :3.000
## NA's :123
## Government Public.Safety
## Mode :logical Mode :logical
## FALSE:1069 FALSE:513
## TRUE :60 TRUE :616
##
##
##
##
## Industry_Role.s.
## Service Provider; Network Operator; :205
## Service Provider; Network Operator; Public Safety; :177
## Service Provider; Network Operator; Equipment Supplier; Public Safety; :175
## Network Operator; : 79
## Service Provider; Network Operator; Public Safety; Property Manager; : 64
## Service Provider; Network Operator; Equipment Supplier; Public Safety; Property Manager; : 62
## (Other) :367

```

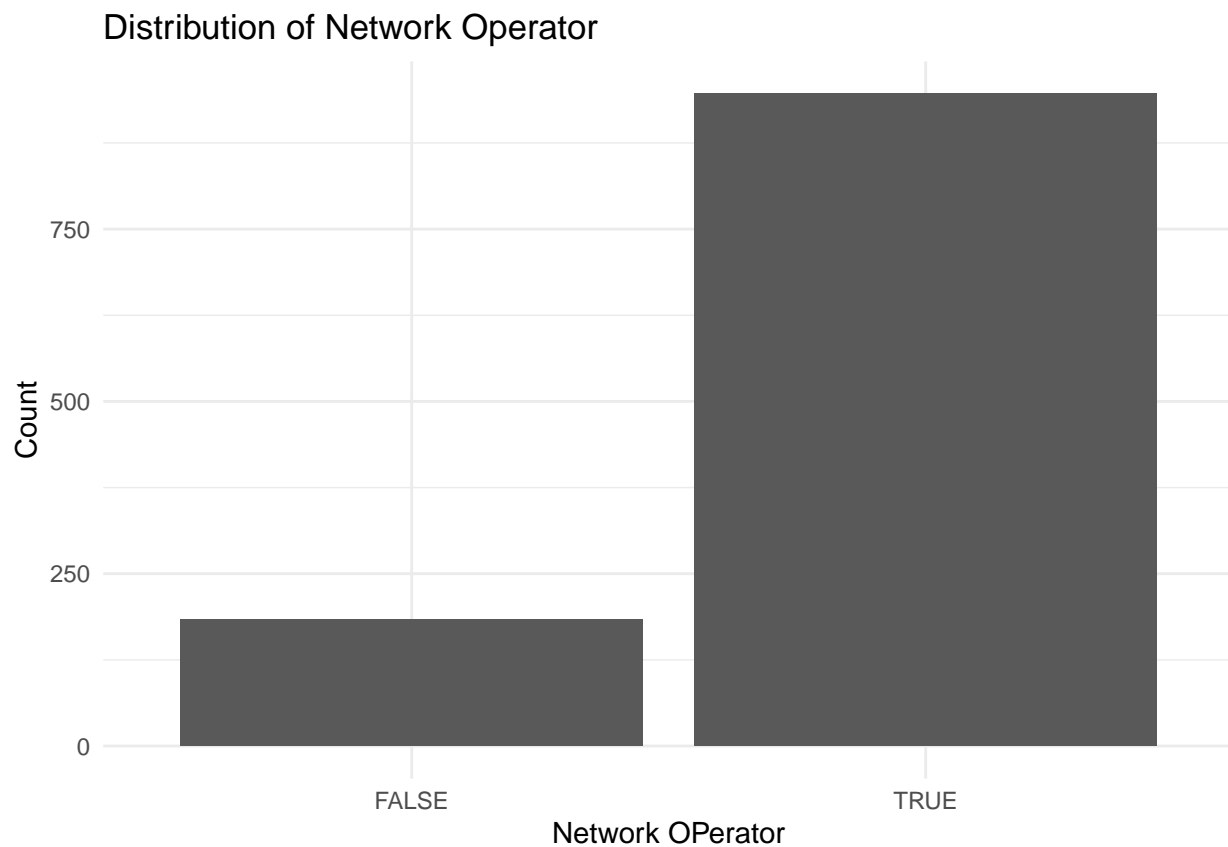
4.2 Distribution of Priority Levels

```
# Plot the distribution of Priority Levels  
ggplot(csric_data, aes(x = Priority)) +  
  geom_bar() +  
  ggtitle("Distribution of Priority Levels") +  
  xlab("Priority Level") +  
  ylab("Count") +  
  theme_minimal()
```



4.3 Network Type Distribution

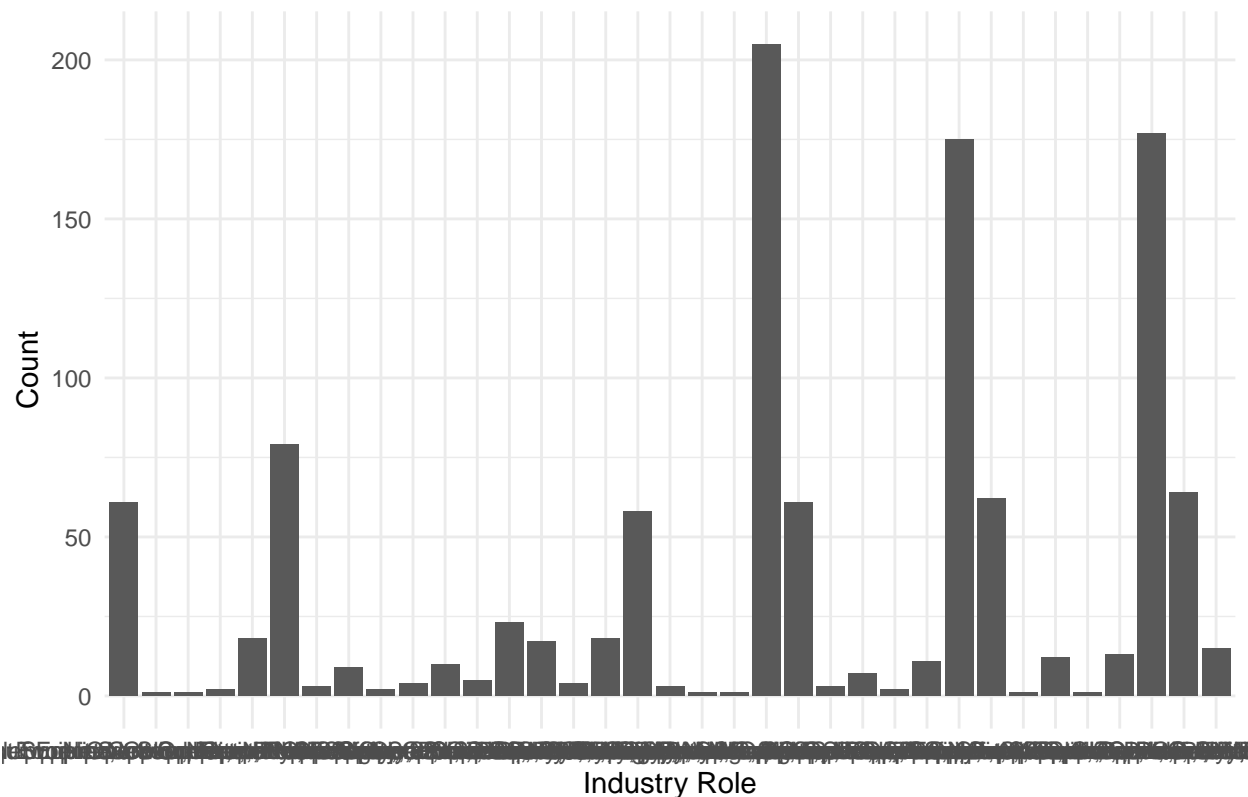
```
# Plot the distribution of Network.Operator  
ggplot(csric_data, aes(x = Network.Operator)) +  
  geom_bar() +  
  ggtitle("Distribution of Network Operator") +  
  xlab("Network OPerator") +  
  ylab("Count") +  
  theme_minimal()
```



4.4 Industry Role Distribution

```
# Plot the distribution of Industry Roles  
ggplot(csric_data, aes(x = Industry.Role.s.)) +  
  geom_bar() +  
  ggtitle("Distribution of Industry Roles") +  
  xlab("Industry Role") +  
  ylab("Count") +  
  theme_minimal()
```

Distribution of Industry Roles



5. Statistical Analysis 5.1 Chi-Squared Test for Priority Level and Network Operator

```
# Chi-squared test for association between Priority Level and Network Operators
table_priority_network <- table(csric_data$Priority, csric_data$Network.Operator)
chi_test_priority_network <- chisq.test(table_priority_network)
chi_test_priority_network
```

```
##
## Pearson's Chi-squared test
##
## data: table_priority_network
## X-squared = 5.0475, df = 3, p-value = 0.1684
```

5.2 Chi-Squared Test for Priority Level and Industry Role

```
# Chi-squared test for association between Priority Level and Industry Role
table_priority_role <- table(csric_data$Priority, csric_data$Industry.Role.s.)
chi_test_priority_role <- chisq.test(table_priority_role)
```

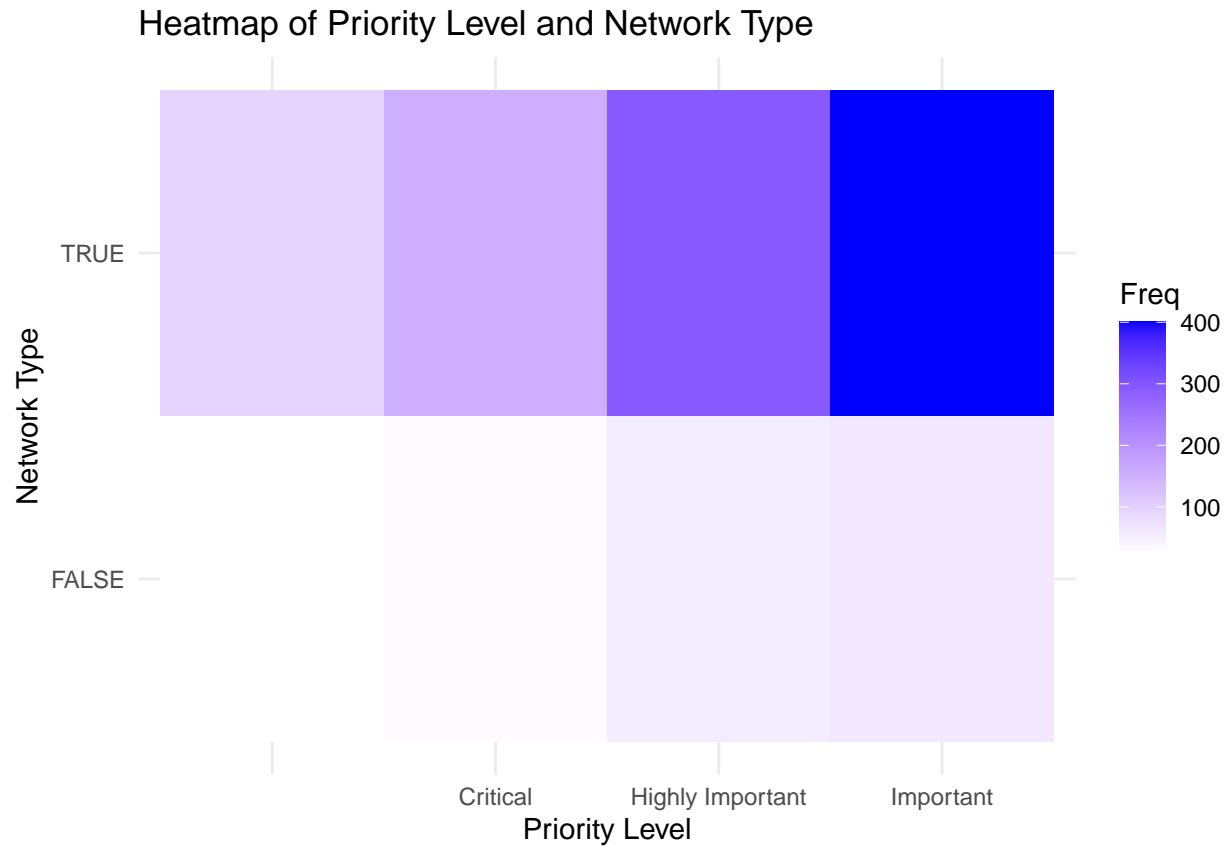
```
## Warning in chisq.test(table_priority_role): Chi-squared approximation may be
## incorrect
```

```
chi_test_priority_role

##
## Pearson's Chi-squared test
##
## data: table_priority_role
## X-squared = 452.06, df = 102, p-value < 2.2e-16
```

6. Visualizing Relationships 6.1 Heatmap of Priority Level and Network Operator Association

```
library(reshape2)
heatmap_data_network <- as.data.frame(table_priority_network)
ggplot(heatmap_data_network, aes(x = Var1, y = Var2, fill = Freq)) +
  geom_tile() +
  scale_fill_gradient(low = "white", high = "blue") +
  labs(title = "Heatmap of Priority Level and Network Type", x = "Priority Level", y = "Network Type") +
  theme_minimal()
```



6.2 Heatmap of Priority Level and Industry Role Association

```
heatmap_data_role <- as.data.frame(table_priority_role)
ggplot(heatmap_data_role, aes(x = Var1, y = Var2, fill = Freq)) +
  geom_tile() +
  scale_fill_gradient(low = "white", high = "blue") +
  labs(title = "Heatmap of Priority Level and Industry Role", x = "Priority Level", y = "Industry Role") +
  theme_minimal()
```




7. Conclusion The analysis of the CSRIC Best Practices dataset provides insights into priority recommendations and the distribution of practices across network operators, industry roles, and public safety aspects. Below are the summary findings:

a) Distribution of Priority Levels

The majority of best practices were found to have lower or medium priority levels, with a smaller subset designated as high priority. High-priority recommendations are predominantly associated with public safety and cybersecurity domains, aligning with the nature of these areas in communications infrastructure.

b) Network Operator and Priority Level Association

An association was observed between network operators and priority levels, for practices linked to mobile and wireless networks. Network operators frequently associated with higher priority recommendations include wireless, mobile, and satellite networks, reflecting the high security and reliability requirements in these areas. Policy advisors should prioritize guidelines for wireless and mobile operators to reinforce security and resilience in these communication infrastructures.

c) Industry Role Distribution

Analysis of industry roles revealed that network operators and service providers are most frequently cited, indicating their central role in implementing best practices. Public safety entities and government roles also play an important part concerning practices relevant to disaster management and emergency response therefore policies should continue to emphasize collaboration between public and private sectors to maintain a robust communication framework.

d) Chi-Squared Tests and Heatmap Visualizations

Chi-squared tests confirmed relationships between priority levels and both network types and industry roles. Heatmap visualizations showed that high-priority best practices are concentrated among network operators and service providers, particularly in roles related to public safety and cybersecurity. Policy implications

include reinforcing best practices for these industry roles to ensure a resilient communications network, particularly in times of crisis or exceptional strain.

e) Recommendations for Policy Advisors

Focus efforts on ensuring compliance and adoption of high-priority recommendations by network operators, particularly in mobile and wireless sectors. Develop targeted guidelines that support the integration of public safety measures across industry roles to enhance the reliability and security of communications infrastructure. Encourage continued collaboration between private network operators and public entities, particularly for roles directly involved in disaster management and emergency response..

8. Reproducibility and Validation The analysis script and cleaned data file are provided for reproducibility. Run this RMarkdown document to replicate the analysis steps.

```
# Save cleaned data for reproducibility (run separately to avoid overwriting)  
write.csv(csric_data, "CSRIC_Best_Practices_Cleaned.csv", row.names = FALSE)
```