

## Introduction:

This document outlines the data structure for our 5G Home Score analysis, which assesses 5G Home WiFi performance by integrating customer-related data from multiple tables (Figure 1).

The goal is to develop a scoring mechanism that combines weighted features to indicate overall WiFi performance."

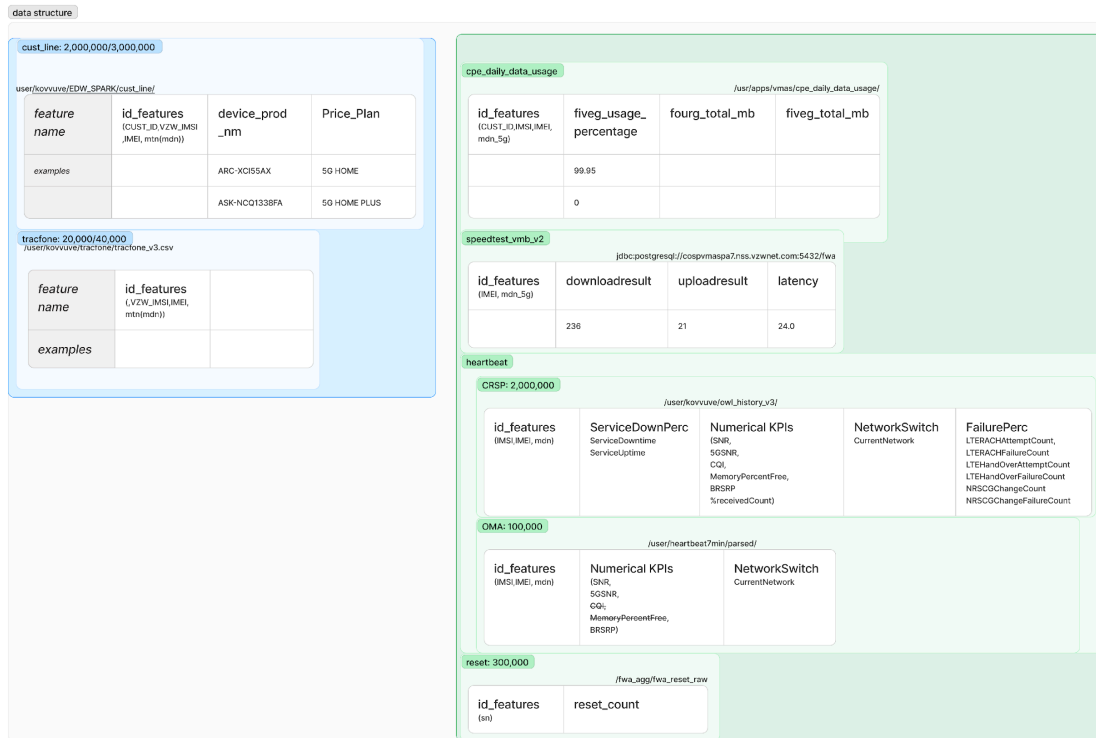


Figure 1 5G home score Profile.

## Data Inventory

### Main Table: cust\_line

#### Description:

This table serves as the primary repository for all 5g\_home\_wifi customers, encompassing essential features such as cust\_id, imsi, imei, and mdn. It also includes other pertinent attributes like price\_plan and device\_product\_name.

## Details:

**device\_product\_name:** Indicates the type of Verizon internet gateway router associated with the customer, such as ASK-NCQ1338FA, FSNO21VA and ARC-XCI55AX.

cpe_model_name	count
ARC-XCI55AX	870396
ASK-NCQ1338FA	521629
WNC-CR200A	257445
FSNO21VA	166187
ASK-NCQ1338	156894
NCQ1338E	51843

**Price\_plan\_description:** Specifies the customer's subscription plan, such as "5G Home" and "5G Home Plus," each with unique characteristics such as speed, video streaming quality, and warranty duration.

**Price\_plan\_code:** further details of price\_plan\_description

The relationship between **device\_product\_name**,

**Price\_plan\_description** and **Price\_plan\_code** are shown

below



cpe model name	Price Plan Description	Price Plan Code	Count
5g Home	ARC-XCI55AX	LTE HOME	50010 138806
		LTE HOME	67577 20997
		LTE HOME PLUS	67584 17565
		5G HOME	50127 274750
		5G HOME	67571 85626
		5G HOME PLUS	67576 200087
		5G HOME PLUS	50129 137201
	ASK-NCQ1338FA	5G HOME	67571 45717
		5G HOME	50127 194123
		5G HOME PLUS	67576 92565
		5G HOME PLUS	50129 114435
		LTE HOME	50010 61507
	WNC-CR200A	5G HOME	67571 67891
		5G HOME PLUS	67576 190968
		LTE HOME PLUS	67584 14808
		LTE HOME	67577 13424
	ASK-NCQ1338	5G HOME	50127 47978
		5G HOME PLUS	50129 21151
		LTE HOME	50010 58316
	FSNO21VA	LTE HOME	38365 29081
		LTE BUSINESS INTERNET 10MBPS 25GB/300GB \$79	48390 18455
		LTE BUSINESS INTERNET 25MBPS 50GB/300GB \$109	48423 20288
		LTE BUSINESS INTERNET 50MBPS 150GB/300GB \$209	48445 10804
		5G BUSINESS INTERNET 100MBPS \$69	51219 61565
		5G BUSINESS INTERNET 200MBPS \$99	53617 44108
		LTE BUSINESS INTERNET PLUS 25MBPS \$119	46799 1126
		LTE BUSINESS INTERNET 10MBPS 25GB/300GB \$79	48390 7339
		LTE BUSINESS INTERNET 25MBPS 50GB/300GB \$109	48423 9446
		LTE BUSINESS INTERNET 50MBPS 150GB/300GB \$209	48445 1705
	NCQ1338E	5G BUSINESS INTERNET 100MBPS \$69	51219 19180
		5G BUSINESS INTERNET 100MBPS \$69	27976 1188
		5G BUSINESS INTERNET 200MBPS \$99	53617 6606

## Tracfone:

In addition to Cust\_Line, roughly 20,000 customers come from prepaid tracfone.

## Data Usage Table:

### cpe\_daily\_data\_usage

#### Description:

This table captures daily data usage metrics, focusing on 5G data consumption. It contains numerical features such as fourg\_total\_mb, fiveg\_total\_mb, and fiveg\_usage\_percentage, reflecting the volume and proportion of data usage across different network generations.

#### Details:

**fourg\_total\_mb:** Volume of 4G data consumed by the customer.

**fiveg\_total\_mb:** Volume of 5G data consumed by the customer.

**data\_usage:** total Volume of data consumed by the customer.

**fiveg\_usage\_percentage:** Percentage of total data usage attributed to 5G connectivity.

## Speed test Table: speedtest\_vmb\_v2

#### Description:

This table records speed-related test results, intermittently conducted for select customers. It includes features such as

downloadresult, uploadresult, and latency, representing download/upload speeds and network latency, respectively.

## Details:

**downloadresult:** Download speed measured during speed tests.

**uploadresult:** Upload speed measured during speed tests.

**latency:** Delay in network communication observed during tests.

Each price plan has related speed threshold

```
vbg_speeds = {  
    '51219': {'Download': 100, 'Upload': 20,  
    'PPLAN_DESC': '5G BUSINESS INTERNET 100MBPS  
$69'},  
    '27976': {'Download': 100, 'Upload': 20,  
    'PPLAN_DESC': '5G BUSINESS INTERNET 100MBPS  
$69'},  
    '53617': {'Download': 200, 'Upload': 10,  
    'PPLAN_DESC': '5G BUSINESS INTERNET 100MBPS  
$99'},  
  
    '48390': {'Download': 10, 'Upload': 6,  
    'PPLAN_DESC': 'LTE BUSINESS INTERNET 10MBPS  
25GB/300GB $79'},  
    '48423': {'Download': 25, 'Upload': 6,  
    'PPLAN_DESC': 'LTE BUSINESS INTERNET 25MBPS  
50GB/300GB $109'},  
    '48445': {'Download': 50, 'Upload': 6,  
    'PPLAN_DESC': 'LTE BUSINESS INTERNET 50MBPS  
150GB/300GB $209'},
```

```
    '46799': {'Download': 25, 'Upload': 6,  
'PPLAN_DESC': 'LTE BUSINESS INTERNET PLUS  
25MBPS $119'},  
    '46798': {'Download': 10, 'Upload': 6,  
'PPLAN_DESC': 'LTE BUSINESS INTERNET PLUS  
10MBPS $89'},  
    }
```

```
vcg_speeds = {  
    '50010': {'Download': 50, 'Upload': 6, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME"},  
    '50011': {'Download': 50, 'Upload': 6, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME"},  
    '67577': {'Download': 25, 'Upload': 4, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME"},  
    '38365': {'Download': 25, 'Upload': 4, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME"}, # Existing entries  
  
    '67584': {'Download': 25, 'Upload': 4, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME PLUS"},  
    '65655': {'Download': 25, 'Upload': 4, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME PLUS"},  
    '65656': {'Download': 25, 'Upload': 4, 'Latency':  
100, 'PPLAN_DESC': "LTE HOME PLUS"},  
  
    '50044': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN_DESC': "5G HOME"},  
    '50055': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN_DESC': "5G HOME"},  
    '50127': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN_DESC': "5G HOME"},
```



'50128': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME"},

'67571': {'Download': 50, 'Upload': 5, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME"},

'67567': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME"},

'50129': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME PLUS"},

'67576': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME PLUS"},

'67568': {'Download': 300, 'Upload': 25,  
'Latency': 30, 'PPLAN\_DESC': "5G HOME PLUS"},

'50116': {'Download': 300, 'Upload': 50,  
'Latency': 30, 'PPLAN\_DESC': "5G HOME PLUS"},

'50117': {'Download': 300, 'Upload': 50,  
'Latency': 30, 'PPLAN\_DESC': "5G HOME PLUS"},

'50130': {'Download': 85, 'Upload': 10, 'Latency':  
30, 'PPLAN\_DESC': "5G HOME PLUS"},

'39425': {'Download': 300, 'Upload': 50,  
'Latency': 30, 'PPLAN\_DESC': "5G HOME INTERNET"},

'39428': {'Download': 300, 'Upload': 50,  
'Latency': 30, 'PPLAN\_DESC': "5G HOME INTERNET"},

}

## Heartbeat Tables: CRSP and OMA

### Description:

These tables monitor the connectivity and performance of 5G Home services, recording activities at frequent intervals (every 5 minutes for CRSP, and every 15 minutes for OMA). These tables contain the majority features used for 5g home Score.

### CRSP:

Features include ServiceDownPercentage, NetworkSwitch, SNR, 5GSR, COI, MemoryPercentFree, BRSRP, 4GRSRP, and failure percentage metrics (LTERACHFailurePerc, LTEHandOverFailurePerc, and NRSCGChangeFailurePerc).

### OMA:

Similar features to CRSP, excluding MemoryPercentFree and failure percentage metrics (LTERACHFailurePerc, LTEHandOverFailurePerc, NRSCGChangeFailurePerc). Both tables provide crucial insights into network stability, performance metrics, and potential issues affecting customer experience.

## Scoring Mechanism:

The scoring mechanism involves:

1. Determining the individual feature scores for each feature, primarily based on their statistical ranking.

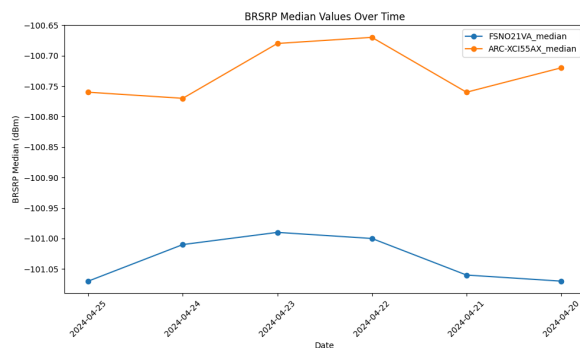
2. Aggregating feature score to Category score.
3. Aggregating Category score to 5g Home score.

## Feature To FeatureScore using distribution statistics

### 1. Monthly distribution:

The statistics (1%, 50%, 99%) comes from previous 1 month distribution, and used to transform current day features into score. The reason is explain below using Example of BRSRP across various device models over several days:

While the daily statistics of BRSRP such as median, 1% percentile, and 99% percentile do not vary drastically day-to-day, there might be exceptions where significant variations occur.



1. To mitigate **the risk of these outliers** influencing the feature scoring unduly, **adopting a one-month** distribution period for calculating these statistics has been considered.
2. Utilizing a month-long aggregation period for statistics will help in creating a more **robust** model, less susceptible to the occasional

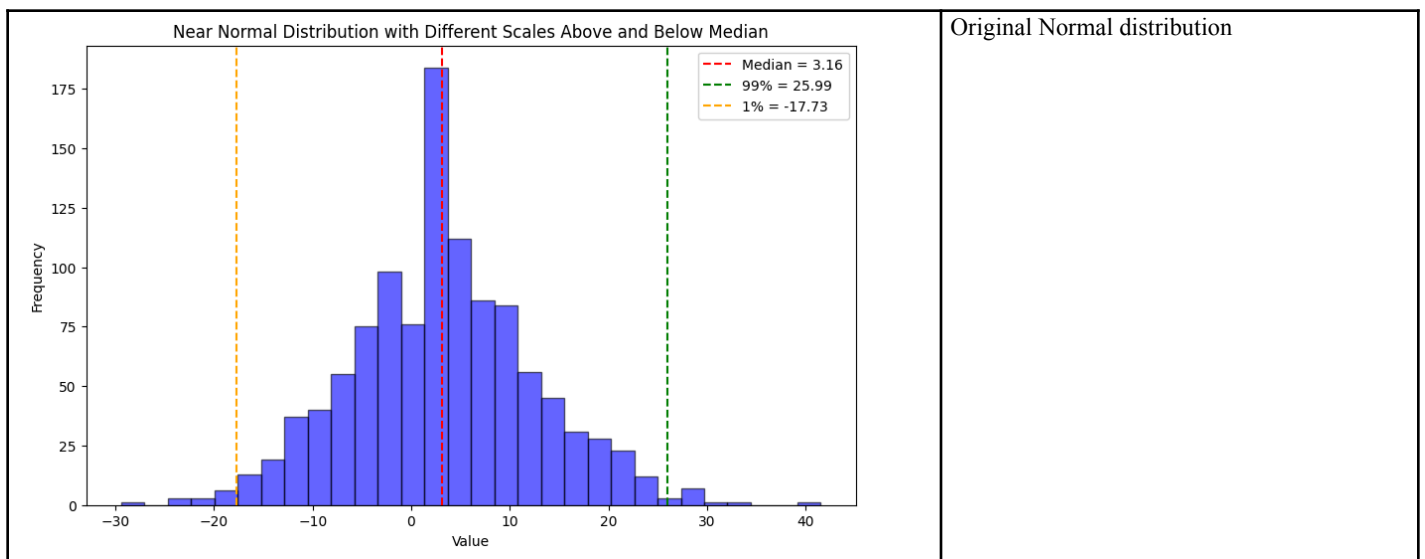
variances that can occur in daily measurements.

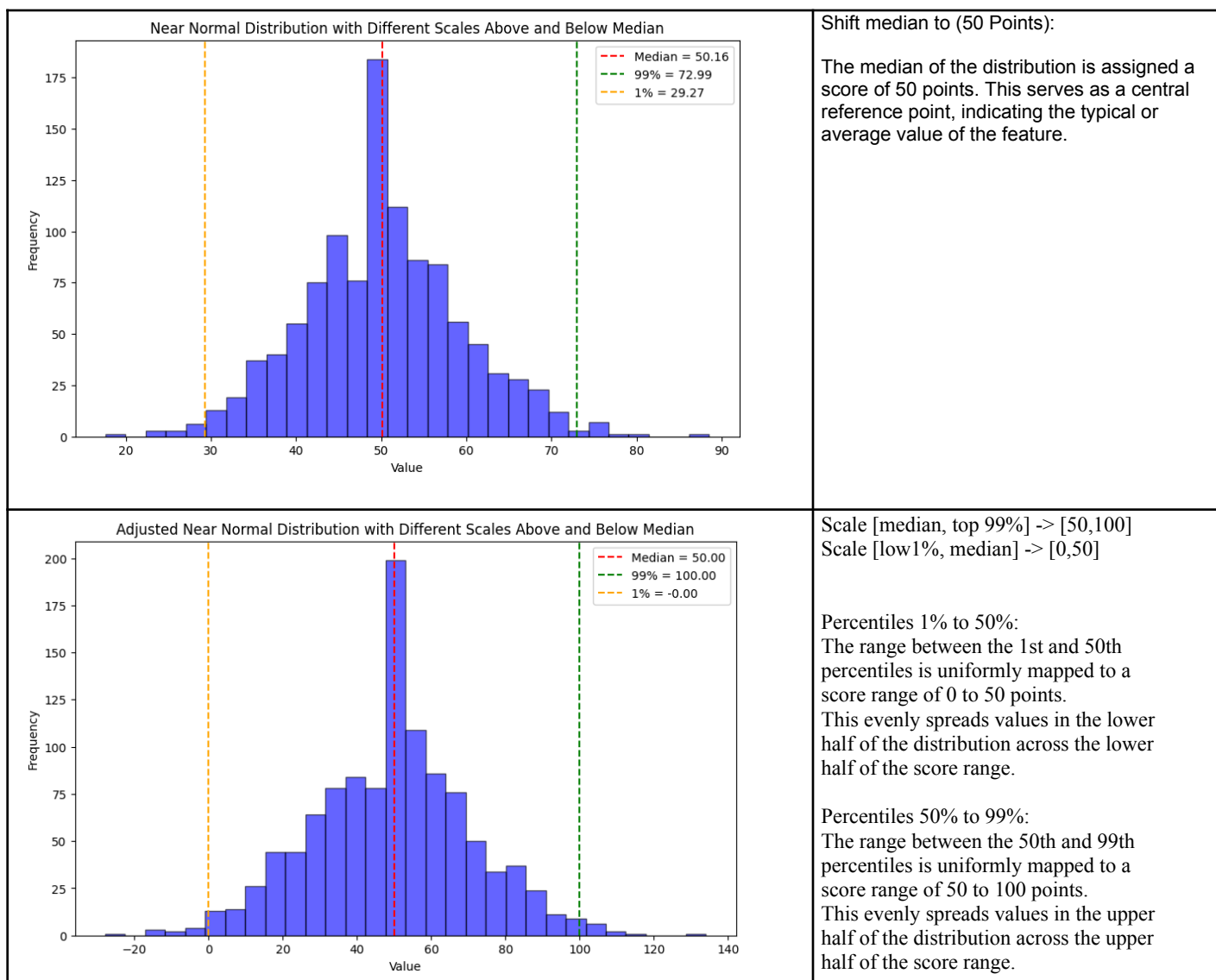
3. The **continuous monitoring** of these metrics will be essential to maintain the accuracy and relevance of the scoring mechanism.

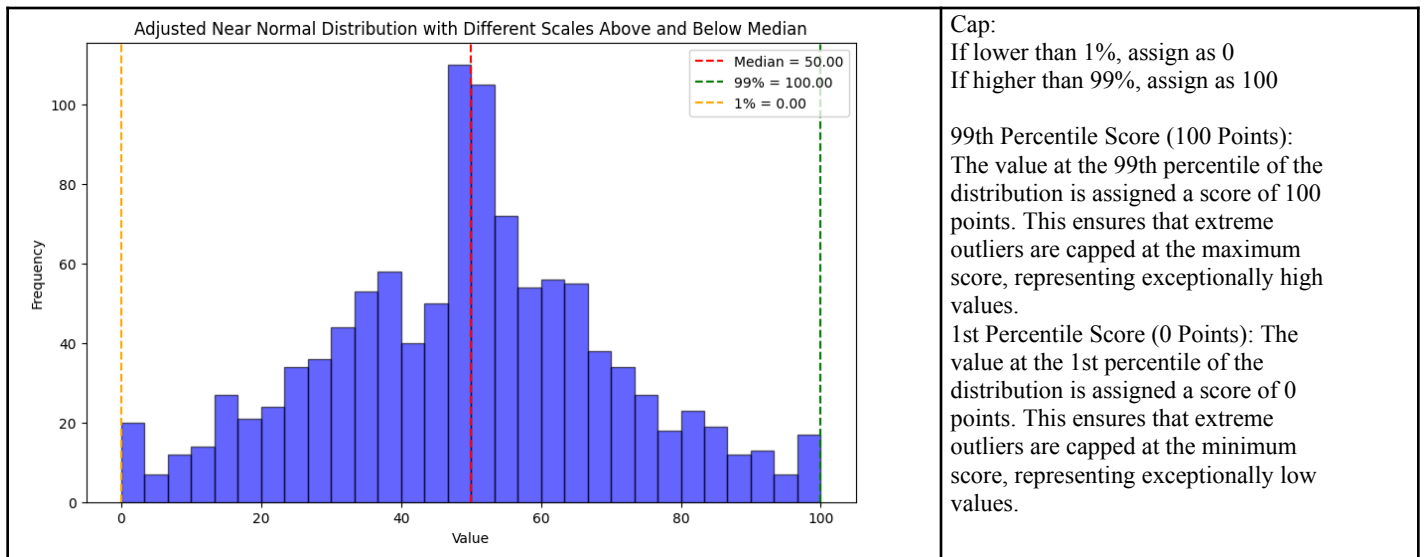
## 2. Normalization and Standardization:

In this section, we outline the scoring mechanism devised to transform a near-normal distribution feature into a standardized score range from 0 to 100. This mechanism aims to provide a meaningful representation of the distribution's characteristics while ensuring that extreme outliers are appropriately handled.

### 1.







### 3. Advantage of standardizing-normalizing over Direct Percentile Scoring:

1. In a normal distribution, most data points cluster around the mean.
2. By standardizing-normalizing, the scores of values near the mean (which represent most of the data) change little relative to each other.
3. This method reduces sensitivity to small changes near the median. Small differences in feature values do not lead to disproportionate score changes.

### 4. Interpretation of Scores:

A score of 10 indicates that the value of the feature falls within the lower half of the distribution but **closer to the extreme lower tail**, suggesting a relatively low value.

A score of 50 indicates that the value of the feature is at the median of the distribution, representing an average value.

A score of 75 indicates that the value of the feature falls within the upper half of the distribution but **not in the extreme upper tail**, suggesting a relatively high value.

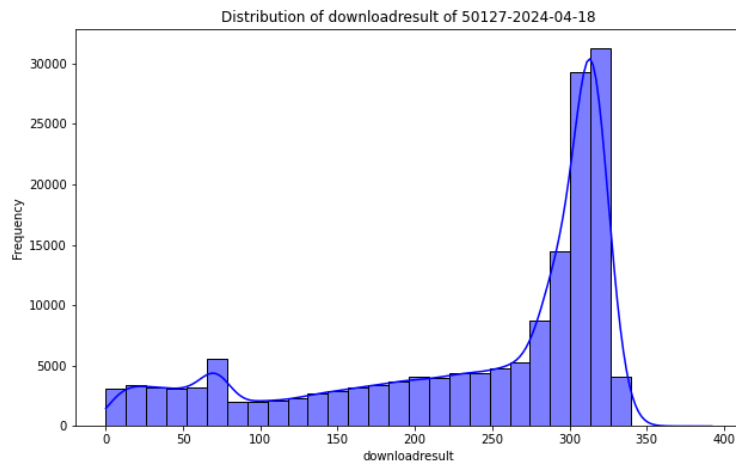
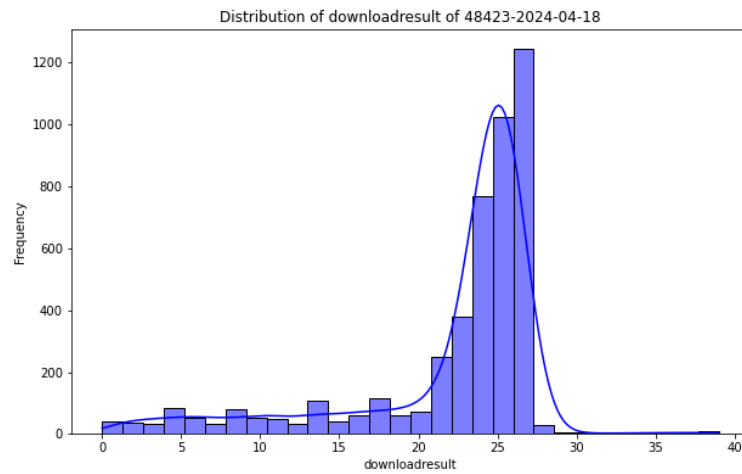
A score of 100 indicates that the value of the feature falls within the highest 1% of the distribution, suggesting an exceptionally high value or extreme outlier.

5. Features follows this method to transform feature to score

```
networkSignalScore = { "4GRSRP", "BRSRP", "CQI", "SNR", "5GSNR" }  
  
dataScore= {  
"5g_uptime", "fiveg_usage_percentage",  
"data_usage" }
```

Feature To FeatureScore using price plan threshold

Instead of using statistics (1%, 50%, 99%) from the feature distribution, this method uses the threshold defined in the price plan advertisement, as shown below.



```
speed_threshold_dict = {
    '51219': [0,100,150],
    '27976': [0,100,150],
    '53617': [0,200,300],

    '48390': [0,10,15],
    '48423': [0,25,40],
    '48445': [0,50,75],
    '46799': [0,25,40],
    '46798': [0,10,15],

    '50010': [0,50,70],
    '50011': [0,50,70],
    '67577': [0,25,40],
    '38365': [0,25,40],

    '67584': [0,25,40],
```



```
'65655': [0,25,40],
'65656': [0,25,40],

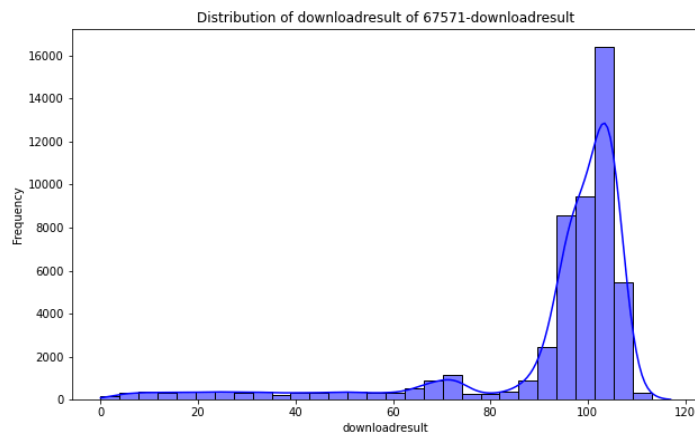
'50044': [0,300,400],
'50055': [0,300,400],
'50127': [0,300,400],
'50128': [0,300,400],
'67571': [0,100,150],
'67567': [0,300,400],

'50129': [0,300,400],
'67576': [0,300,400],
'67568': [0,1000,1500],
'50116': [0,1000,1500],
'50117': [0,1000,1500],
'50130': [0,300,400],

'39425': [0,1000,1500],
'39428': [0,1000,1500],
}
```

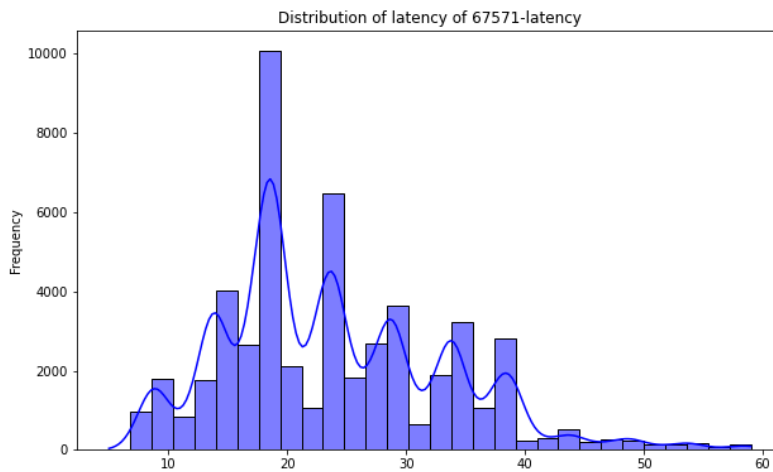
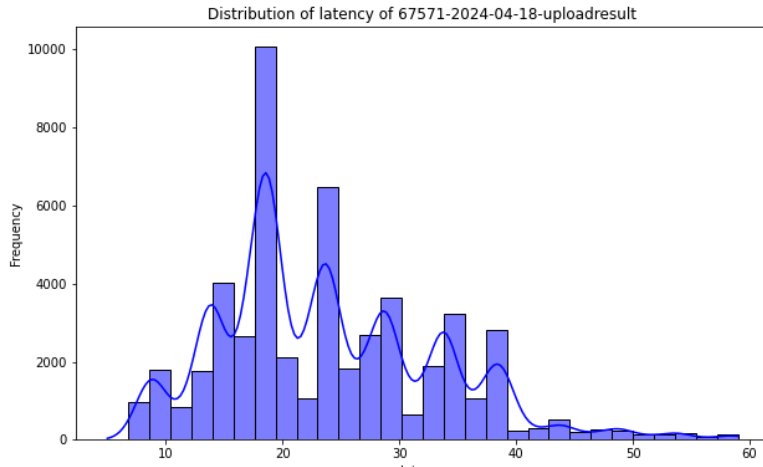
## 1. Interpretation:

**Score 50 means it reaches the advertised speed tire.**



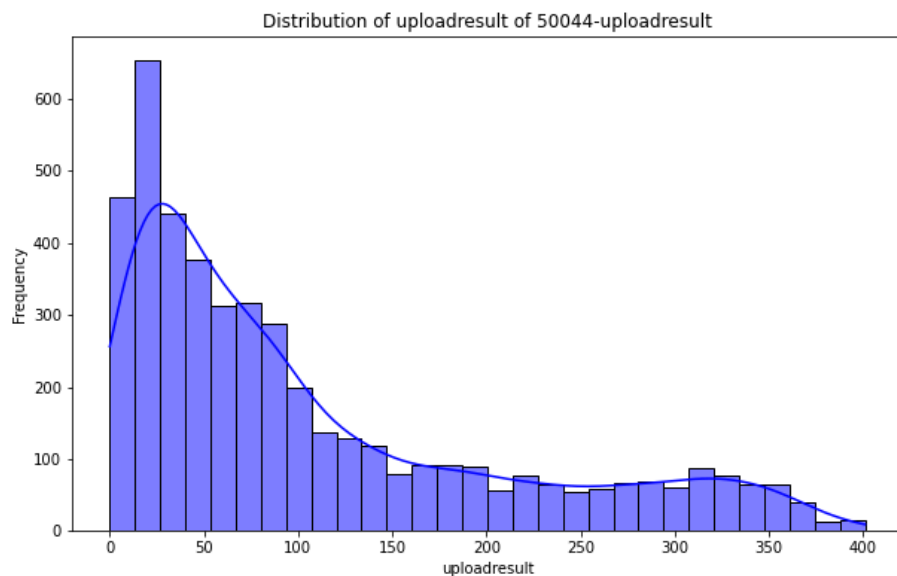
New/Existing Plan ID (examples)	67571 - for new customers only
Spectrum to be utilized on Price Plan	n77 + LTE
APN	V5GA01INTERNET
Max Up To Speed - DL	100Mbps
Max Up To Speed (Peak) - UL	10Mbps =>11Mbps
Customer disclosure - DL	50 - 85Mbps DL w/ speeds up to 100Mbps
Customer disclosure - UL	5-10Mbps
Customer disclosure - Latency (RTT)	30ms
Data Cap/ Threshold - DL / UL	unlimited

2. Features follows this method to transform feature to score



	mmWave - basic	mmWave - plus	mmWave - 5GHome (basic)	mmWave - 5GHome +
FWA			Price Plan 1	Price Plan 2
Launch Date	2022	2022	2023 (new)	2023 (new)
New/Existing Plan ID (examples)	50044 Consumer 50055 Employee	50116 Consumer 50117 Employee	50044 Consumer 50055 Employee	50116 Consumer 50117 Employee
Spectrum to be utilized on Price Plan	mmWave + LTE	mmWave + LTE	mmWave + LTE fallback	mmWave + LTE fallback
APN		V5GA01INTERNET	V5GA01INTERNET	V5GA01INTERNET

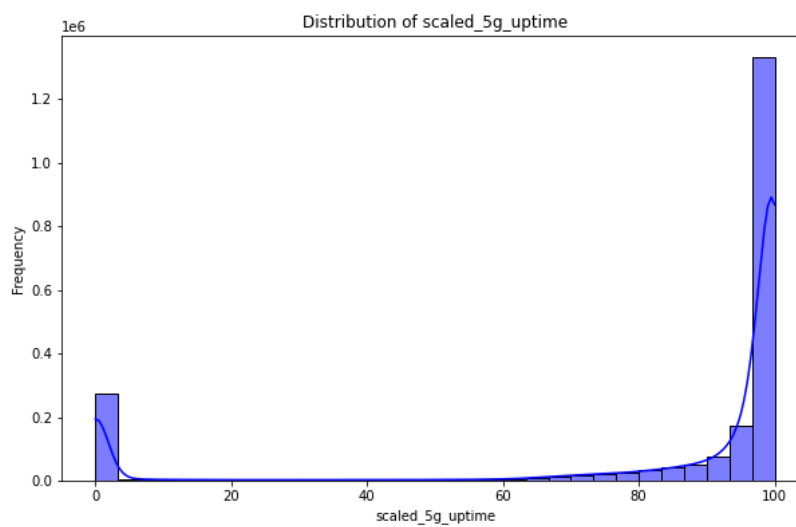
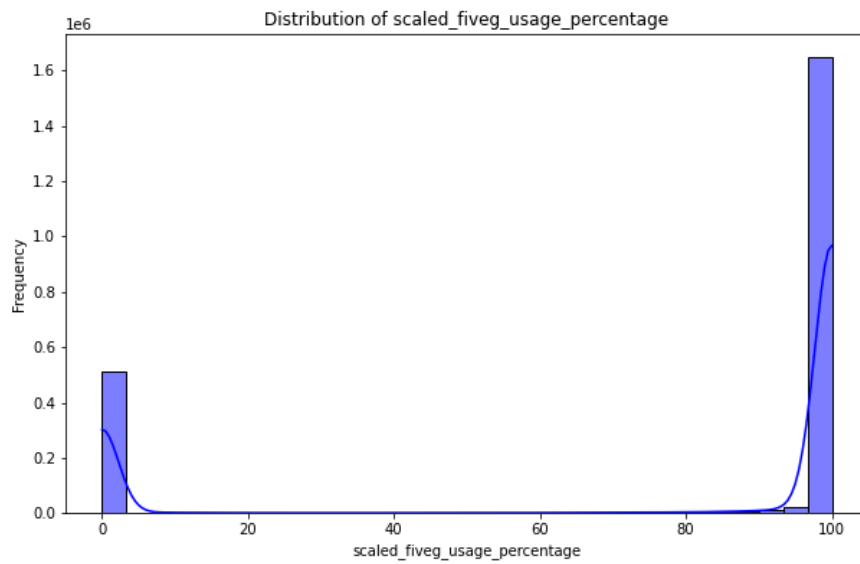
<b>Max Up To Speed - DL</b>	no cap	no cap	400Mbps	1.5Gbps
<b>Max Up To Speed (Peak) - UL</b>	no cap	no cap	20Mbps	75Mbps
<b>Customer disclosure - DL</b>	85-300Mbps	300-1000Mbps	85-300Mbps	300-1000Mbps
<b>Customer disclosure - UL</b>	<b>typical 10Mbps</b>	typical 50Mbps	10-20Mbps	25-75Mbps
<b>Customer disclosure - Latency (RTT)</b>	30ms	30ms	30ms	30ms



## Feature To FeatureScore with Special

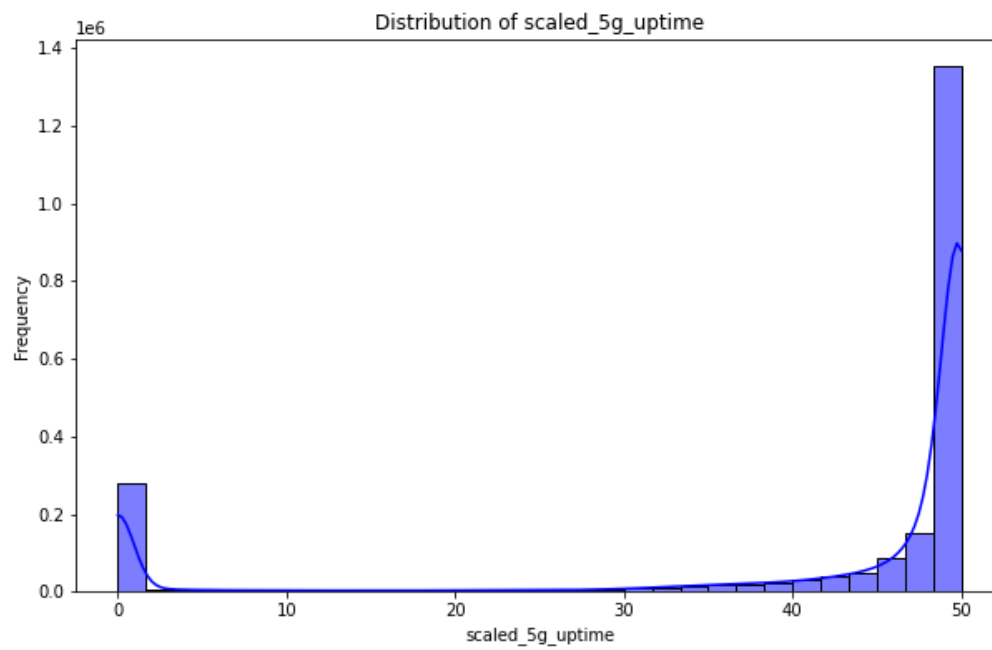
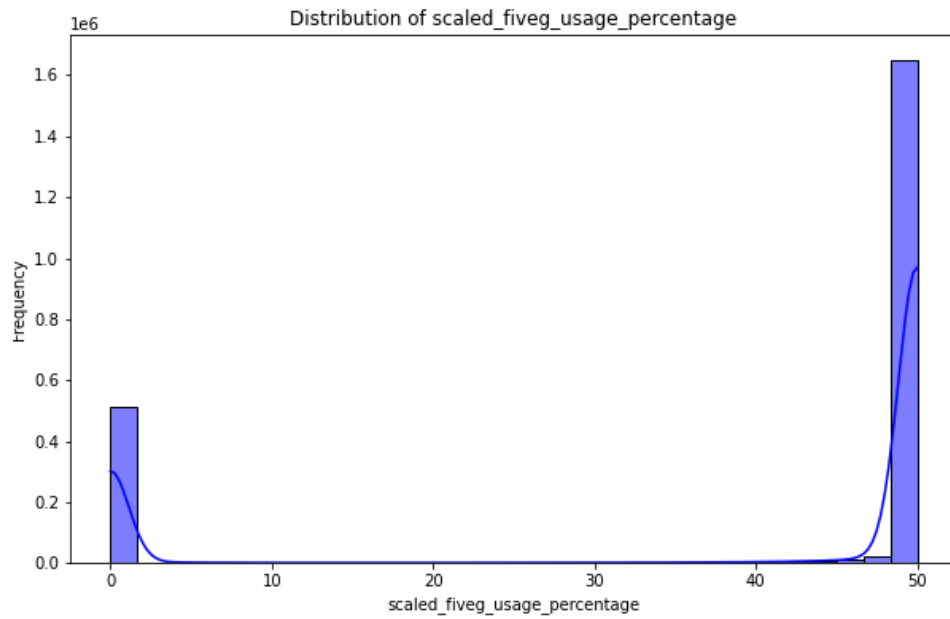
### Case:

1. feature--(function)-->feature\_score
2. Function should be reasonable
3. The feature\_score should be consistent interpretable: below 50 is bad



feature-(function)-->feature\_score, feature\_to\_score function should lead to consistent interpretation of score.

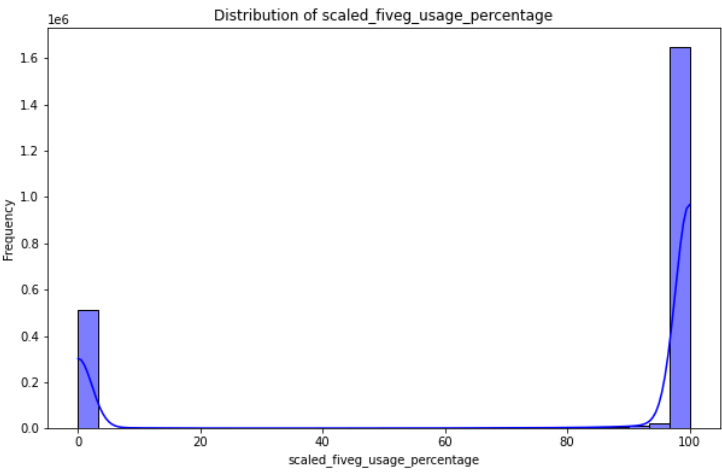
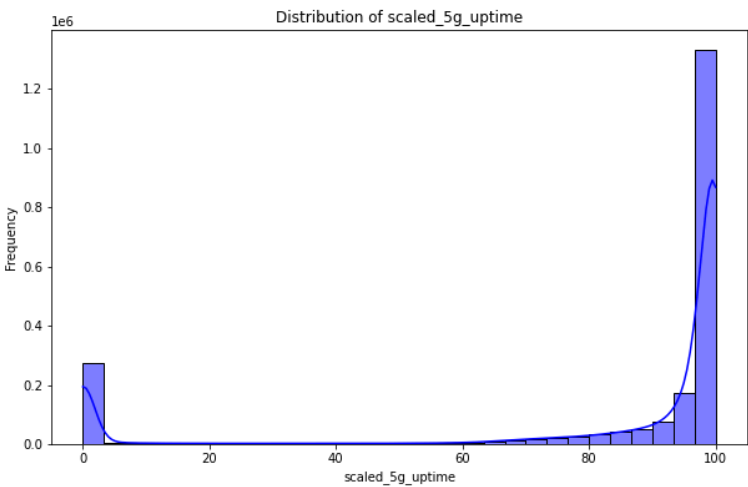
Case 1: Fiveg\_usage\_percentage, 5g\_uptime

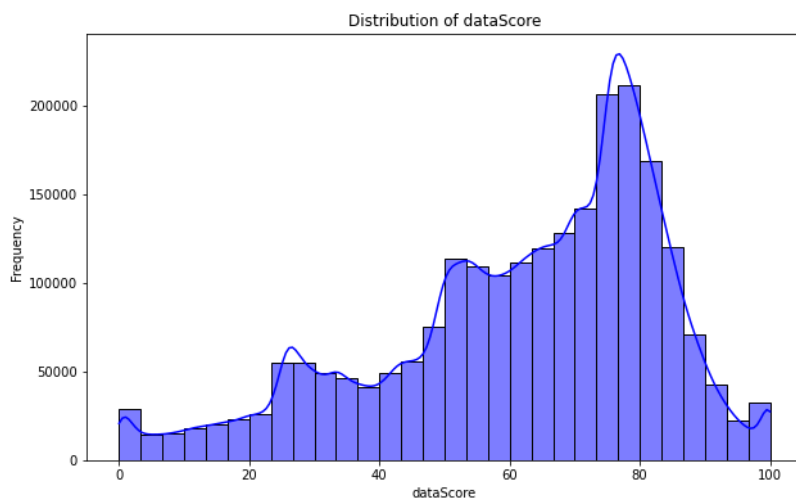
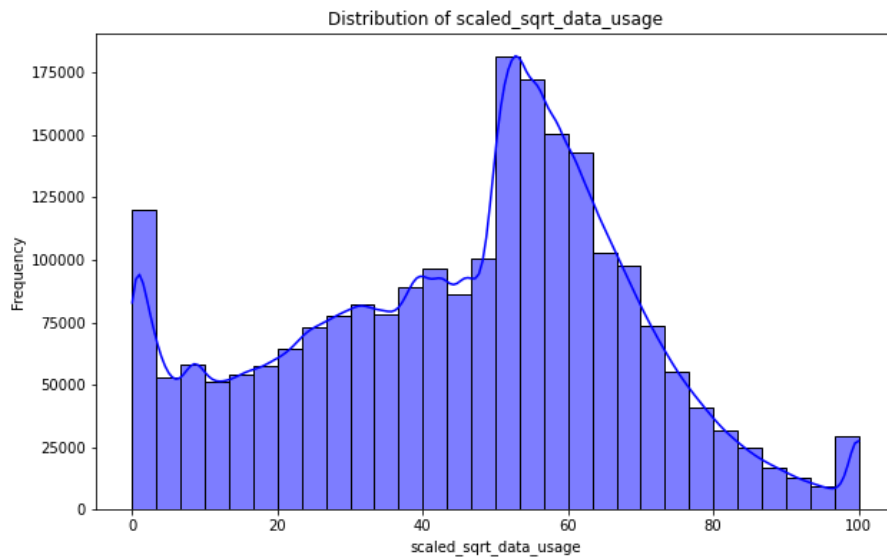


**benefit:**

```
dataScore_weights = {
    "scaled_5g_uptime": 0.25,
    "scaled_fiveg_usage_percentage": 0.25,
    "scaled_sqrt_data_usage": 0.5,
}
```

If use 100, instead of 50

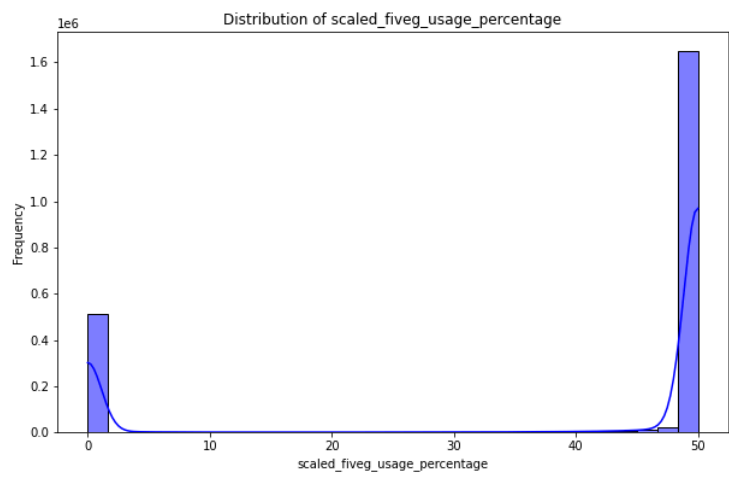
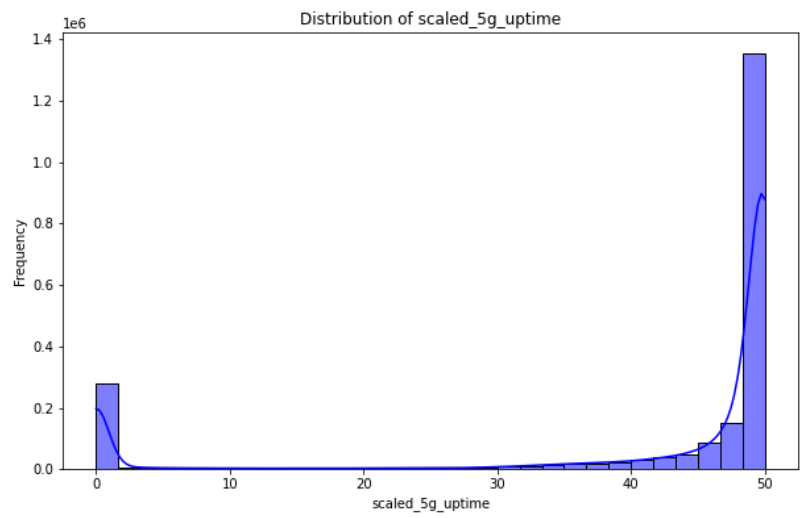


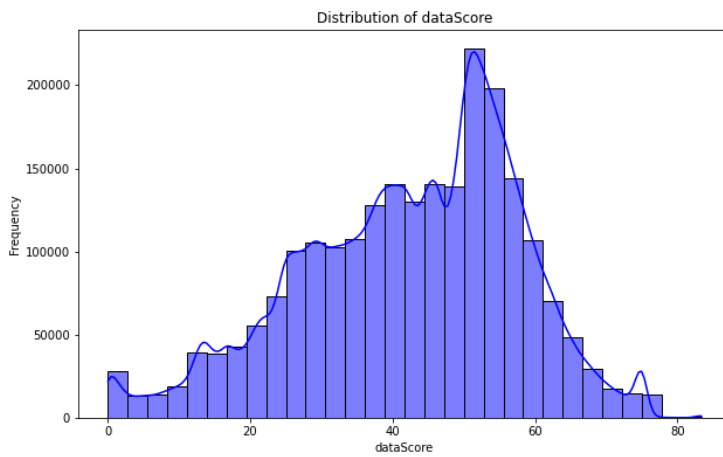
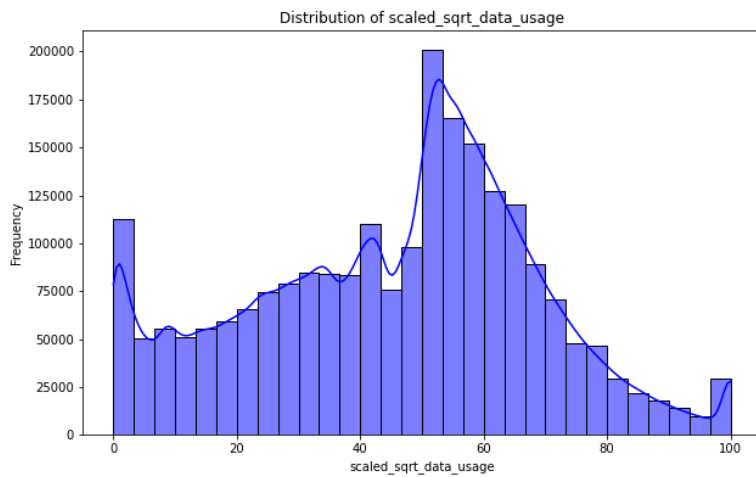


if you perform badly on data usage, not represented in category score, you may end up with 70.

**If use 50**





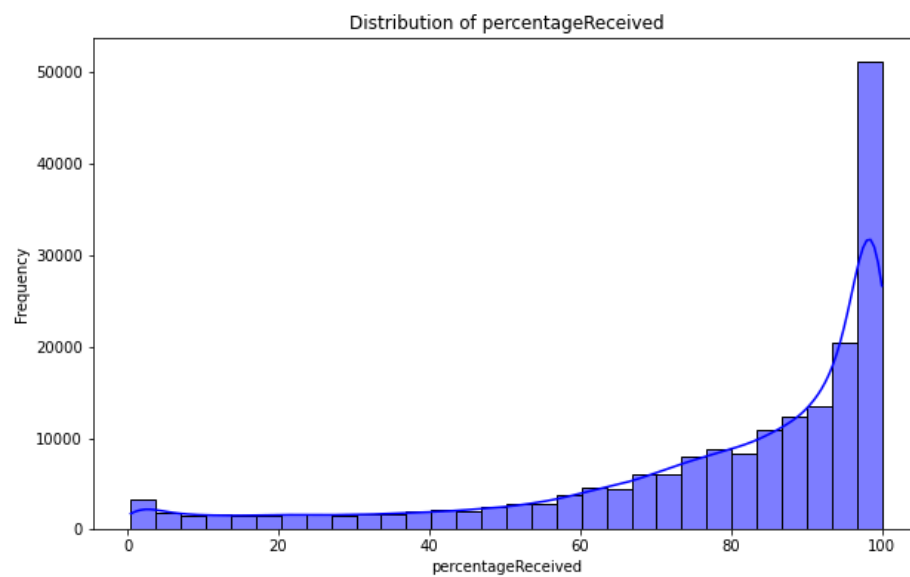
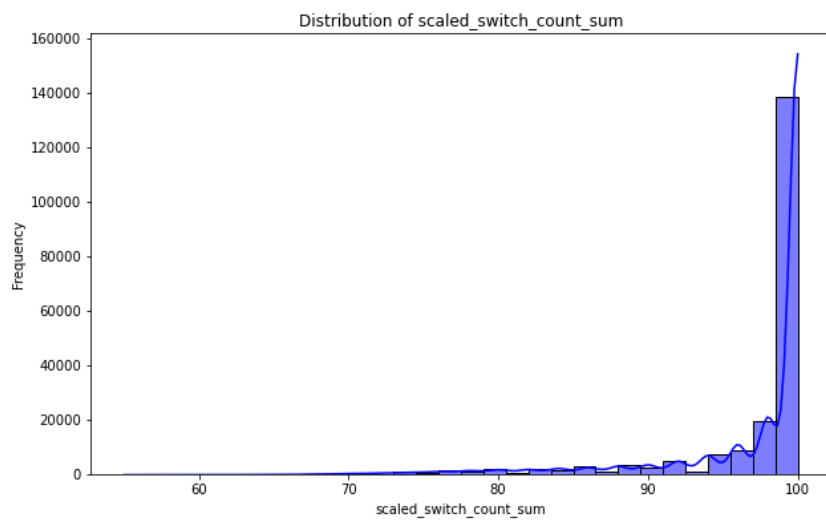


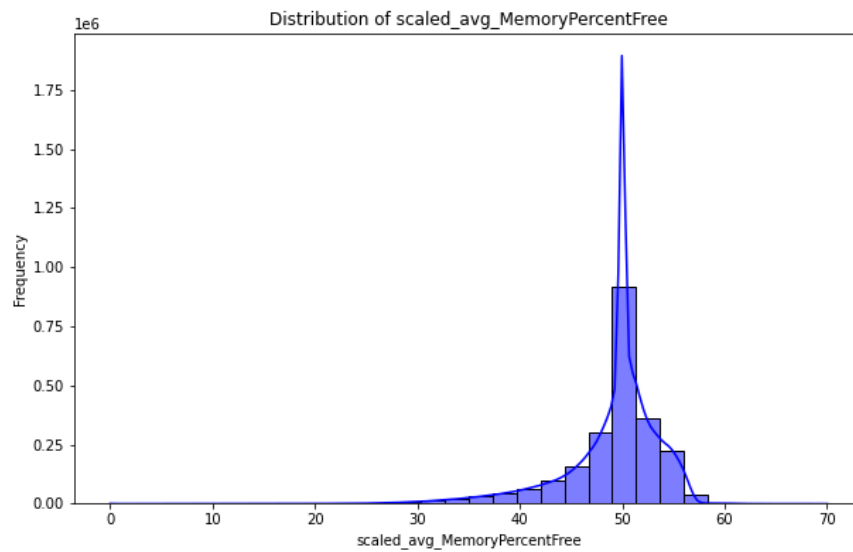
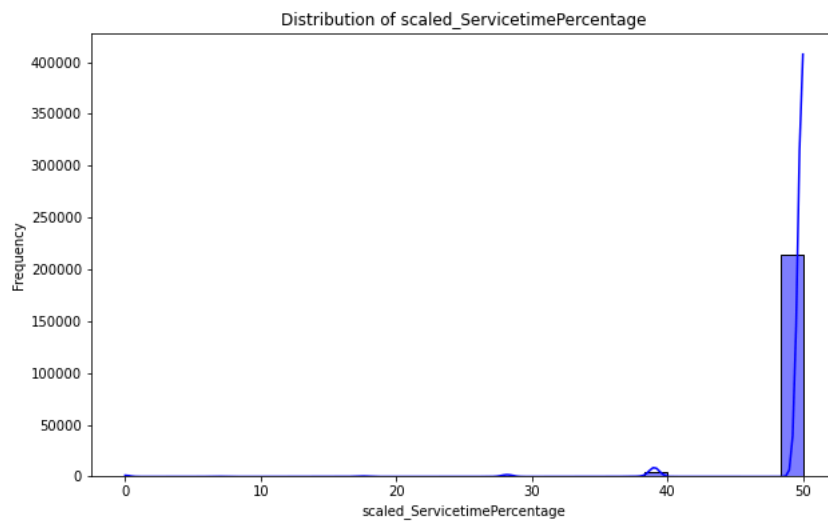
\*map [50,80] to [50,100]

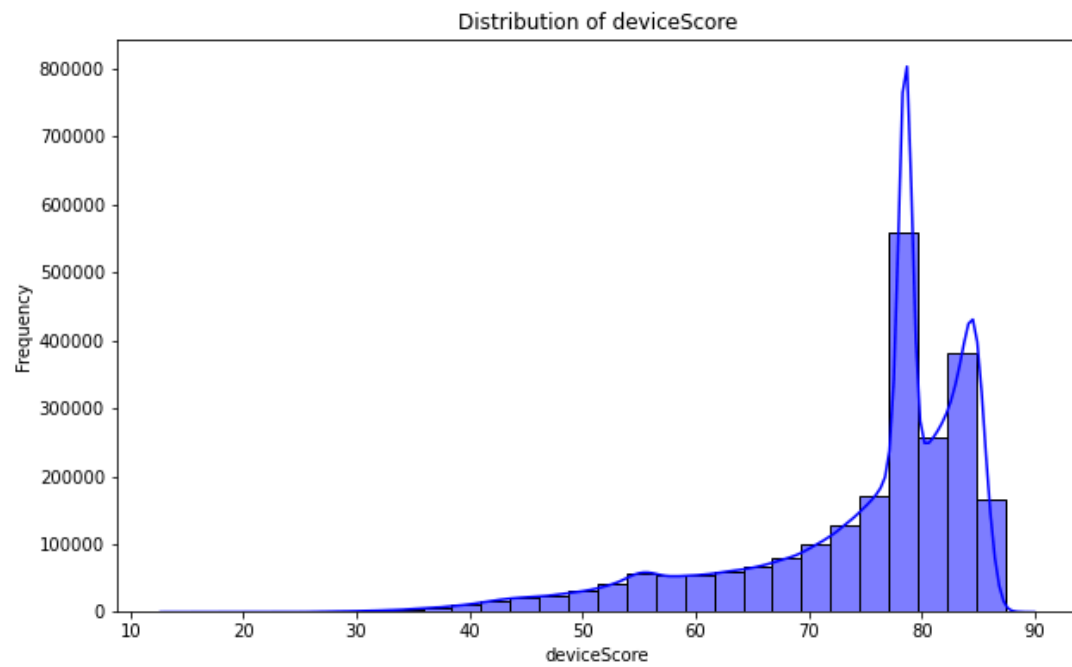
1. Focus on low score, if you perform badly, no matter which score, always represented in category score
2. Not overshadow other feature effect

Case 2: switch\_count\_sum, percentageReceived

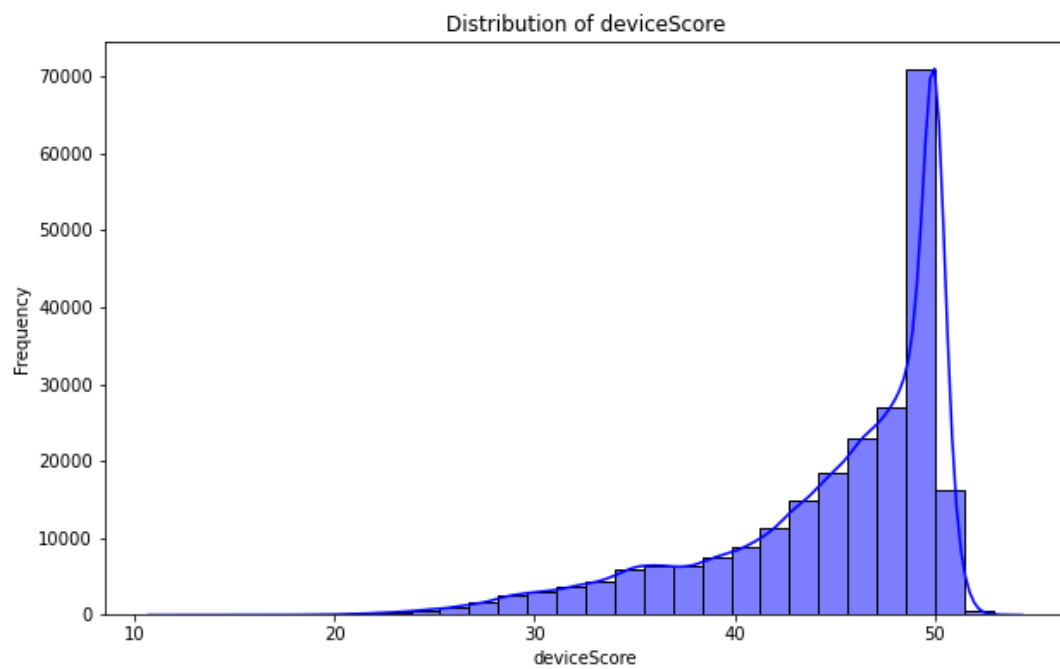
```
deviceScore_weights = {
    "scaled_switch_count_sum": 0.3,
    "scaled_percentageReceived": 0.3,
    "scaled_reset_count": 0.3,
    "avg_MemoryPercentFree": 0.1
}
```



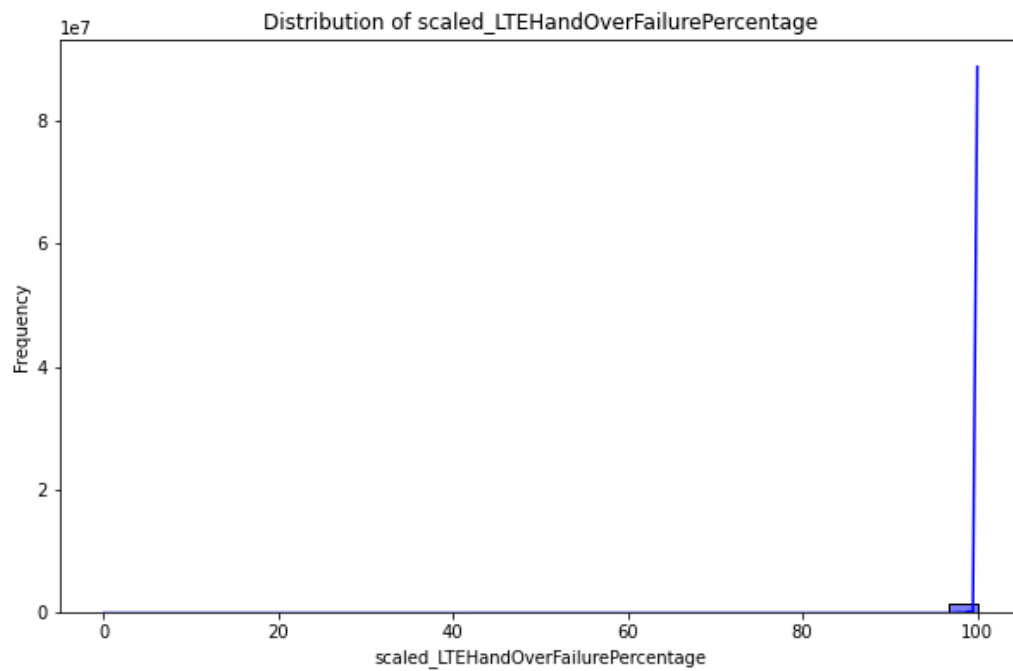




After switch [0,100] to [0,50]

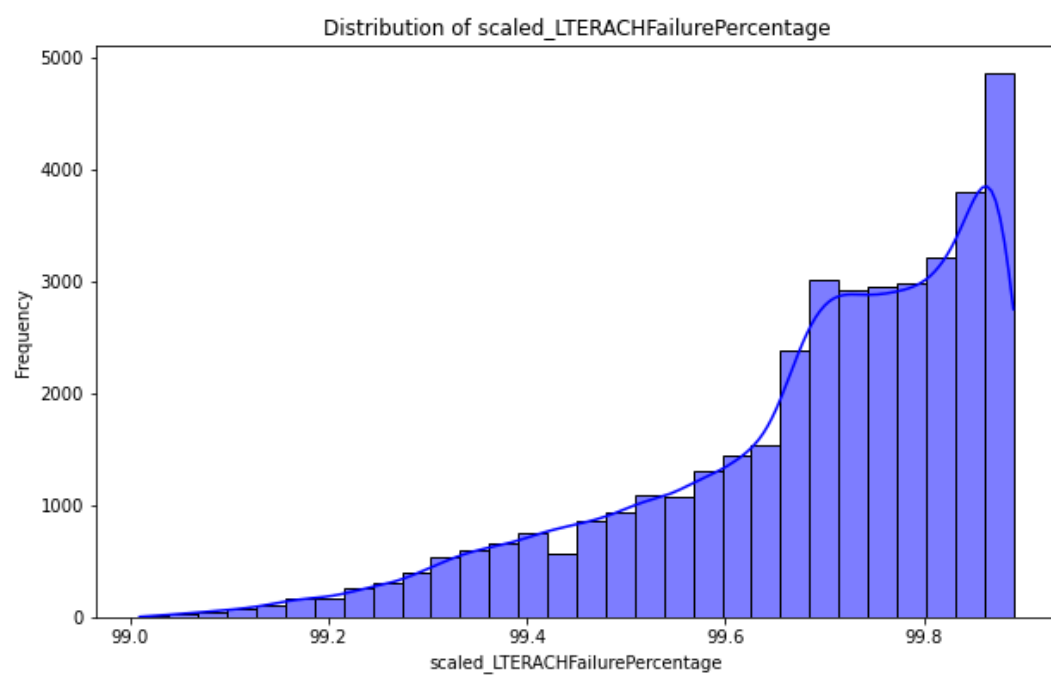


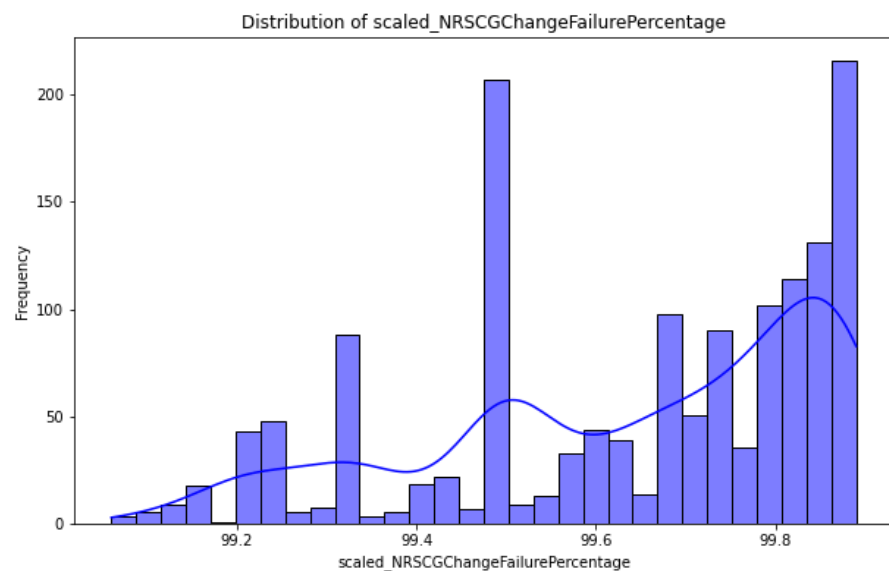
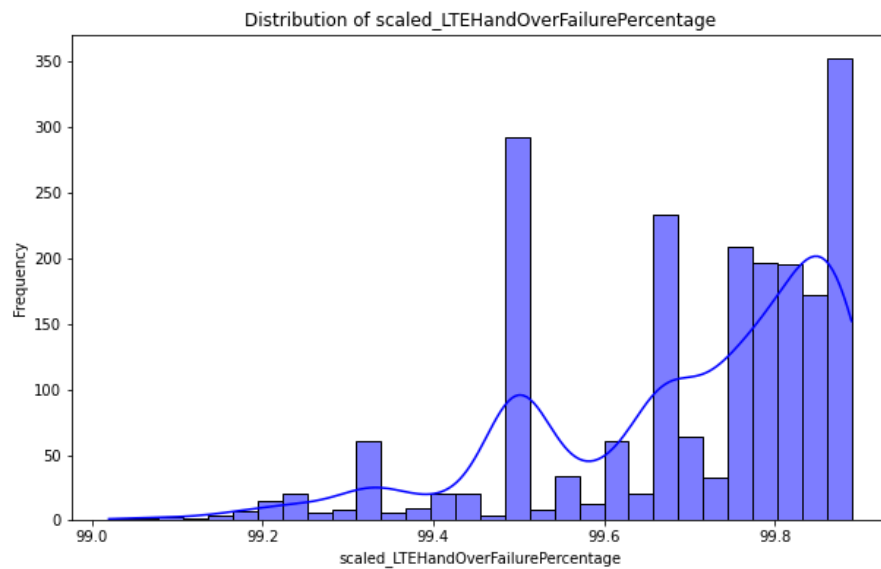
Case 3: LTERACHFailurePercentage, LTEHandOverFailurePercentage, NRSCGChangeFailurePercentage, reset\_count, ServicetimePercentage



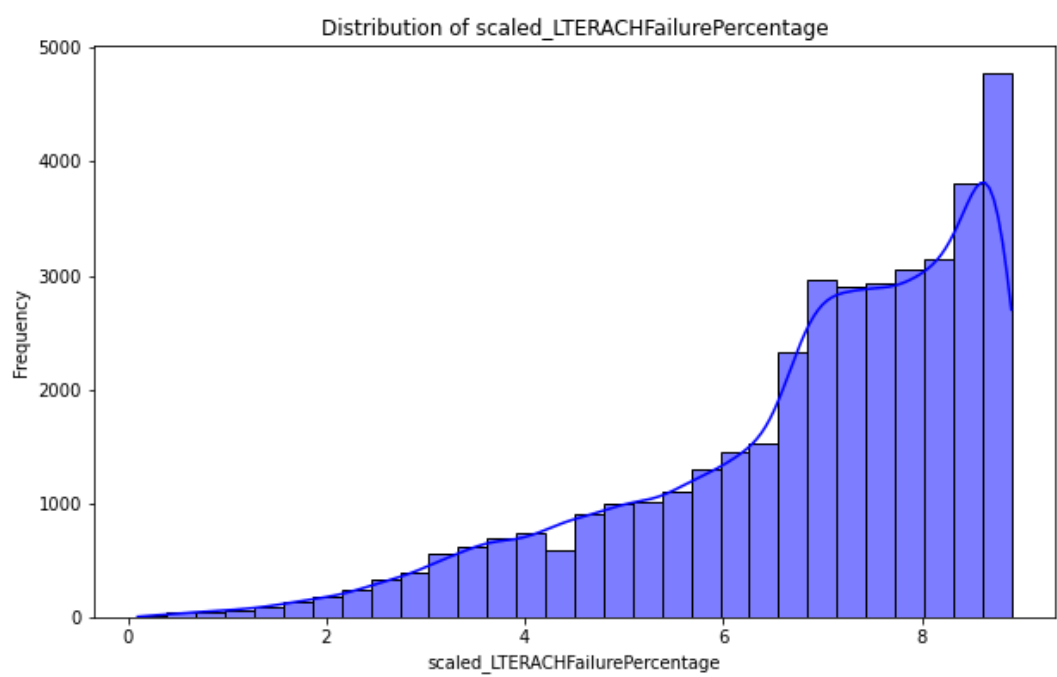
>99.9: 1782002

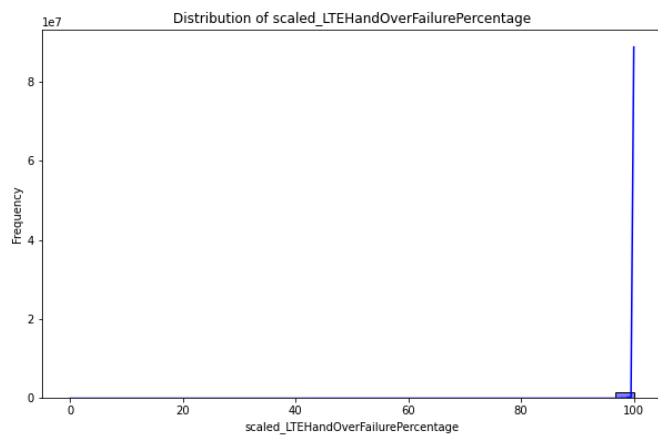
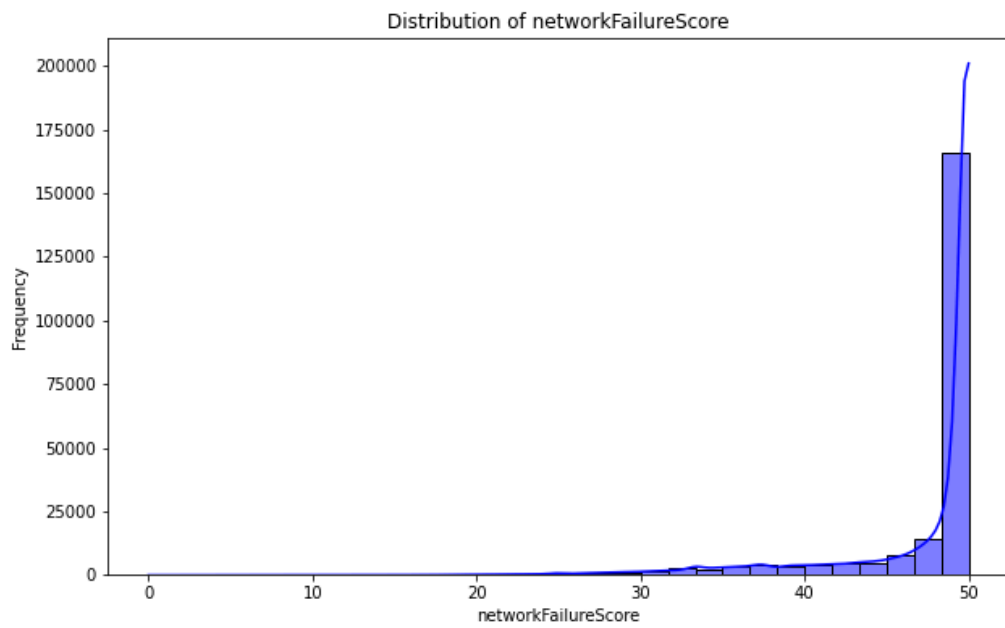
<99.9: 392541



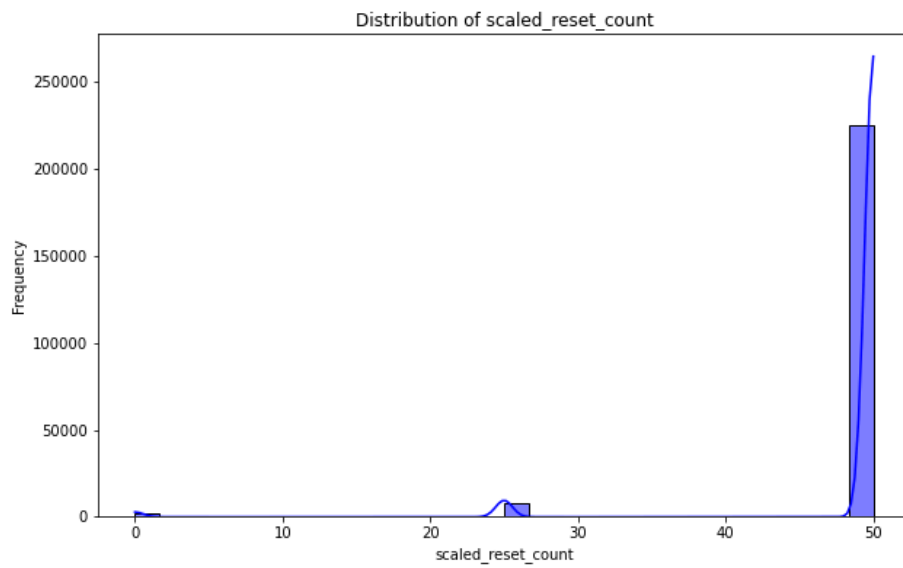
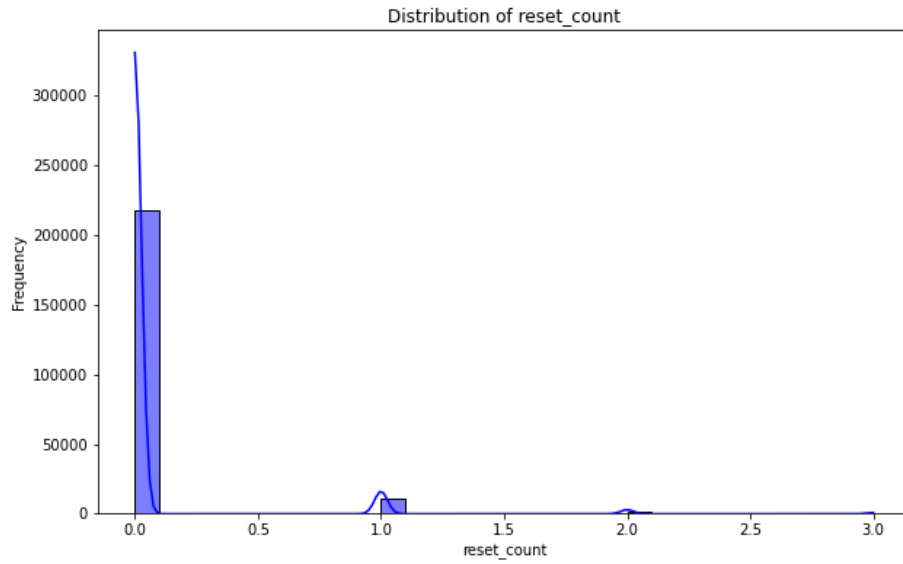








reset\_count



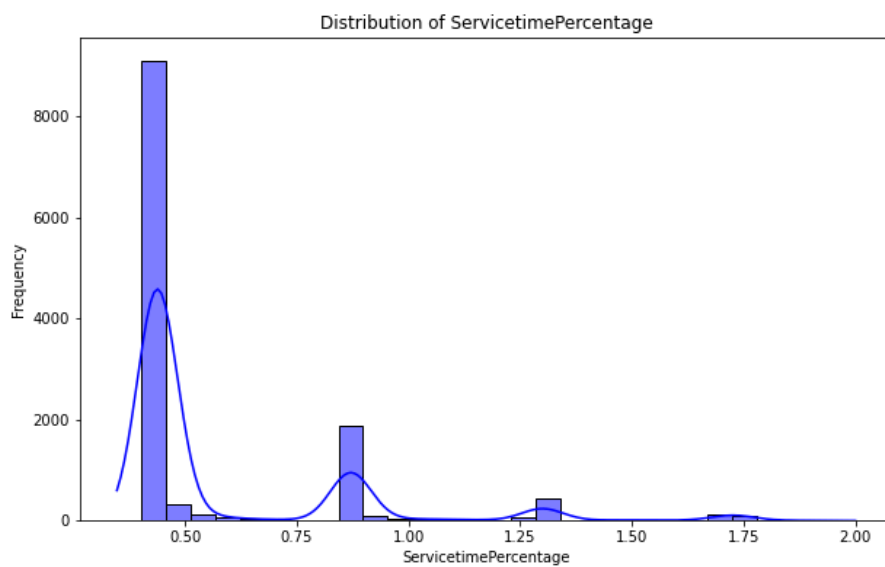
```
print( df_failure.filter( col("reset_count")==3 ).count() )  
print( df_failure.filter( col("reset_count")==2 ).count() )
```

```
5576  
19134
```

```
print( df_failure.filter( col("reset_count")==1 ).count() )  
print( df_failure.filter( col("reset_count")==0 ).count() )
```

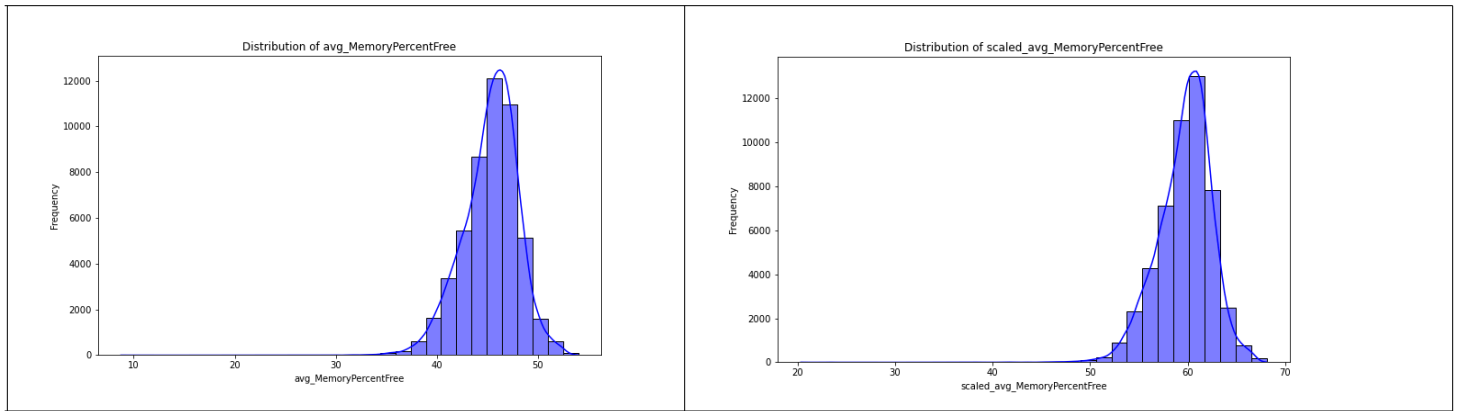
```
109142  
2170636
```

ServicetimePercentage  
[0,2]->[50,0]

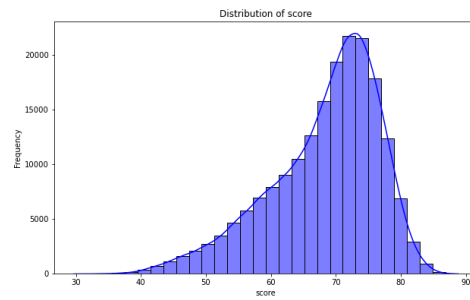


Shift

avg\_MemoryPercentFree

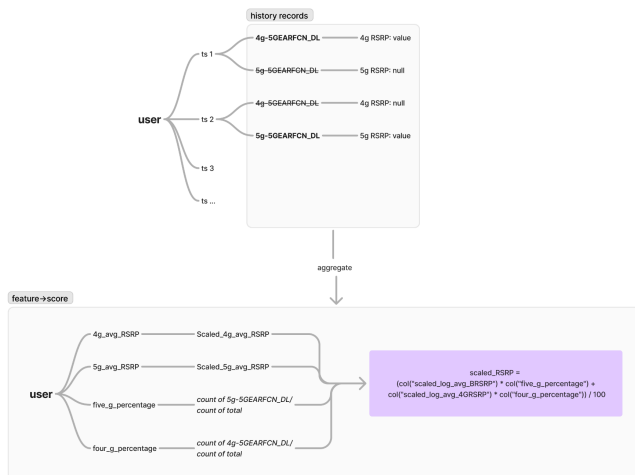
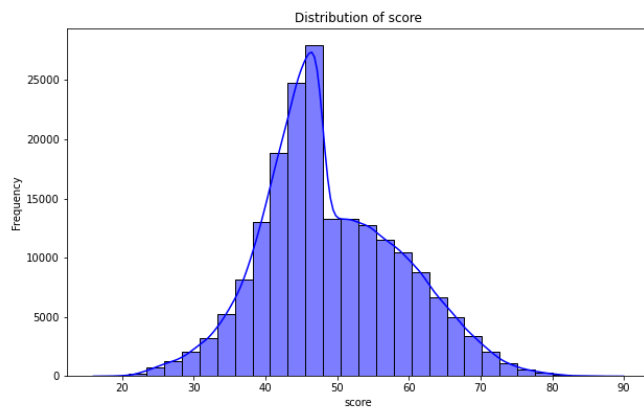
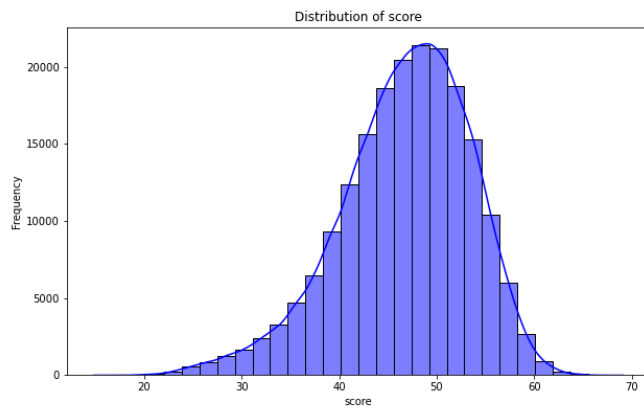


**[0,100]**



**[0,50]:** below 50ish is bad(below average), and is consistent over all feature\_score, categorical score, overall\_score

Drawback, Above 50, different interpretation



## Feature aggregation

### 1. Feature Score to Category Score

$$S_i = \frac{\sum_{j=1}^{n_i} \begin{cases} w_{ij} \cdot f_{ij} & \text{if } f_{ij} \text{ is not null} \\ 0 & \text{if } f_{ij} \text{ is null} \end{cases}}{\sum_{j=1}^{n_i} \begin{cases} w_{ij} & \text{if } f_{ij} \text{ is not null} \\ 0 & \text{if } f_{ij} \text{ is null} \end{cases}}$$

### 2. Category Score to 5g Home Score

$$T = \frac{\sum_{i=1}^3 \begin{cases} w_i \cdot S_i & \text{if } S_i \text{ is not null} \\ 0 & \text{if } S_i \text{ is null} \end{cases}}{\sum_{i=1}^3 \begin{cases} w_i & \text{if } S_i \text{ is not null} \\ 0 & \text{if } S_i \text{ is null} \end{cases}}$$

```
networkSpeedScore_weights = {  
    "scaled_uploadresult": 0.2,  
    "scaled_downloadresult": 0.2,  
    "scaled_latency": 0.2  
}  
networkSignalScore_weights = {  
    "scaled_log_avg_4GRSRP": 0.2,  
    "scaled_log_avg_BRSRP": 0.2,  
    "scaled_avg_CQI": 0.2,  
    "scaled_log_avg_SNR": 0.2,  
    "scaled_log_avg_5GSRP": 0.2,  
}  
networkFailureScore_weights = {  
  
    "scaled_LTERACHFailurePercentage": 0.5,  
  
    "scaled_LTEHandOverFailurePercentage": 0.25,  
  
    "scaled_NRSCGChangeFailurePercentage": 0.25,  
}
```

```

dataScore_weights = {
    "scaled_5g_uptime": 0.2,
    "scaled_fiveg_usage_percentage": 0.5,
    "scaled_sqrt_data_usage": 0.3,
}

deviceScore_weights = {
    "scaled_switch_count_sum": 0.25,
    "scaled_percentageReceived":0.25,
    "scaled_reset_count":0.25,
    "avg_MemoryPercentFree":0.25
}

```

## Appendix:

cpe_model_name	2024-04-25 BRSRP_median	2024-04-25 BRSRP_lower_5_percentile	2024-04-25 BRSRP_top_95_percentile
FSNO21VA	-101.07	-118.39	-68.89
ARC-XCI55AX	-100.76	-119.21	-70.55
ASK-NCQ1338FA	-100.2	-118.13	-71.02
WNC-CR200A	-99	-117.7	-69.66
tracfone	-102.74	-121.08	-71.38
NCQ1338E	-101.14	-119	-69.49
ASK-NCQ1338	-100.37	-119.82	-70.23
Others	-82.52339	-113.01	-49.563793
	2024-04-24 BRSRP_median	2024-04-24 BRSRP_lower_5_percentile	2024-04-24 BRSRP_top_95_percentile
FSNO21VA	-101.01	-118.4	-68.9
ARC-XCI55AX	-100.77	-119.19	-70.57



ASK-NCQ1338FA	-100.19	-118.08	-71
WNC-CR200A	-98.98	-117.68	-69.61
tracfone	-102.63	-121.21	-71.26
NCQ1338E	-101.11	-119	-69.26
ASK-NCQ1338	-100.38	-119.83	-70.17
Others	-82.57234	-113.12718	-49.492977
	2024-04-23 BRSRP_median	2024-04-23 BRSRP_lower_5_percentile	2024-04-23 BRSRP_top_95_percentile
FSNO21VA	-100.99	-118.4	-69
ARC-XCI55AX	-100.68	-119.12	-70.51
ASK-NCQ1338FA	-100.12	-118.03	-71
WNC-CR200A	-98.85	-117.58	-69.52
tracfone	-102.52	-121.13	-71.33
NCQ1338E	-101.086586	-119	-69.12
ASK-NCQ1338	-100.26	-119.76	-70.25
Others	-82.477	-112.95	-49.59368
	2024-04-22 BRSRP_median	2024-04-22 BRSRP_lower_5_percentile	2024-04-22 BRSRP_top_95_percentile
FSNO21VA	-101	-118.3	-68.87
ARC-XCI55AX	-100.67	-119.06	-70.5
ASK-NCQ1338FA	-100.14	-118.02	-70.95
WNC-CR200A	-98.87	-117.58	-69.44
tracfone	-102.5	-121.24	-71.17
NCQ1338E	-101.06	-119	-69.32
ASK-NCQ1338	-100.25	-119.64	-70.14

Others	-82.636696	-113.03008	-49.623066
	2024-04-21 BRSRP_median	2024-04-21 BRSRP_lower_5_percentile	2024-04-21 BRSRP_top_95_percentile
FSNO21VA	-101.06	-118.38	-68.9
ARC-XCI55AX	-100.76	-119.12	-70.43
ASK-NCQ1338FA	-100.21	-118.05	-70.94
WNC-CR200A	-98.93	-117.71	-69.46
tracfone	-102.6	-121.44	-71.58
NCQ1338E	-101	-119	-69.31
ASK-NCQ1338	-100.21	-119.63	-70
Others	-82.48329	-113	-49.7276
	2024-04-20 BRSRP_median	2024-04-20 BRSRP_lower_5_percentile	2024-04-20 BRSRP_top_95_percentile
FSNO21VA	-101.07	-118.28	-68.85
ARC-XCI55AX	-100.72	-119.03	-70.42
ASK-NCQ1338FA	-100.16	-118.02	-70.97
WNC-CR200A	-98.88	-117.58	-69.52
tracfone	-102.52	-121.31	-71.18
NCQ1338E	-101.1	-119	-69.46
ASK-NCQ1338	-100.24	-119.71	-70.1
Others	-82.444565	-112.84	-49.721775

Covariance:

