**Data Collection**

This section gives us more insight on how data was collected from Dataset 1 and Dataset 2.

Data Collection of Dataset -1

Dataset 1 was sourced from “https://www.kaggle.com/datasets/prasertk/internet-broadband-and-mobile-speeds-by-country,” which contains data on mobile internet speeds across various countries. The original dataset includes six columns, featuring broadband speed, latency, upload and download mobile speeds, and the number of devices tested. For simplicity, we selected only two rows representing the highest number of devices tested, which is suitable for our descriptive analysis.

Approach of data set 1 collection.

The choice of this dataset on internet speeds is based on several important factors. First, internet connectivity is essential in today’s world, impacting areas such as remote work, online education, and entertainment. Analysing variations in internet speeds across different regions and times can offer valuable insights for consumers, businesses, and policymakers. Second, the dataset provides a thorough overview of performance metrics, including download and upload speeds and latency, all of which are vital for evaluating internet service quality. Moreover, the data collection methodology—utilizing real-time tests from a reliable platform—ensures both accuracy and credibility. Finally, this dataset enables trend analysis over time, helping stakeholders make informed decisions about infrastructure investments and service enhancements, ultimately improving connectivity for all users.

Briefly about dataset collected by Ookla

Ookla is a technology company renowned for its widely used internet speed testing service, Speed test. Established in 2006, it offers tools for users to measure their internet connection speed, including download and upload rates and latency. With a global user base, Ookla provides insights into network performance for both consumers and businesses. In addition to speed tests, the company gathers and analyses data on internet performance trends, delivering valuable reports on global broadband speeds and connectivity. Ookla's mission is to empower users with knowledge about their internet services and improve the overall online experience.

Briefly about data speed in Mbps

Data speed, measured in megabits per second (Mbps), represents the rate at which data is transmitted over a network. It reflects how quickly information can be downloaded from or uploaded to the internet. For instance, a speed of 25 Mbps means that 25 megabits of data can be transferred each second. Generally, higher Mbps values result in faster website loading times, smoother video streaming, and improved performance for online gaming and video conferencing. Various factors, including network congestion, distance from the service provider, and the type of connection (such as fibre, DSL, or cable), can affect the actual data speeds users experience.

Data Collection of Dataset -2

Dataset 2 was gathered from an Android smartphone by recording the notification history log from 9 AM to 5 PM on a specific day (September 9, 2024). The raw dataset contains a single row that details the various times each notification was sent by the Android system to the user.

Approach of data set 2 collections.

The dataset was collected on a specific day by capturing the notification history through Android logs and filtering the notifications for a defined time frame to create a more manageable dataset. This dataset consists of 125-time rows in hh:mm:ss format. Analysing the timestamps of notifications can yield valuable insights for several reasons. For example, researchers can identify patterns in user engagement, including peak times when users are most likely to interact with their devices.

Briefly about choice of data set -2

This analysis can inform strategies for optimizing notification timing to boost user engagement. Additionally, understanding when notifications are sent helps evaluate their effectiveness; for instance, notifications delivered during off-peak hours may see lower engagement compared to those sent when users are more active. Also, in light of growing concerns about digital well-being, analysing notification timing can shed light on their impact on users’ mental health and productivity. Insights into the frequency and timing of notifications can guide better practices for managing digital interruptions. Overall, analysing when notifications are sent can yield actionable insights that enhance user experience, engagement, and overall app effectiveness.

**Descriptive Analysis**

Dataset-1 : Internet Speeds of various counties in the world for the year 2022 (Measured in Mbps).

|  |  |
| --- | --- |
| Measurement | Value (Measured in Mbps). |
| Sample Mean | 49.04 |
| Sample Median | 33.59863 |
| Max Internet Speed | 242.77 |
| Min Internet Speed | 5.45 |
| Standard deviation | 41.66207 |
| Sample range | 237.32 |
| Quartile 1 | 21.82129 |
| Quartile 2 | 33.70068 |
| Quartile 3 | 61.66699 |
| Geomean | 36.95869 |
| Coefficient of variation | 0.849554 |
| Mode | NA |

From the measures presented in the table, we can interpret the following statistics:

* **Sample Mean**: The average internet speed worldwide is 49.04 Mbps (megabits per second).
* **Sample Median**: The median, or central value, is 33.59 Mbps (megabits per second).
* **Sample Range**: The sample range is 237.2 Mbps (megabits per second).
* **Sample Standard Deviation**: The standard deviation relative to the mean is 41.66 Mbps (megabits per second).
* **First Quartile (Q1)**: The first quartile represents the median of the lower half of the dataset, indicating that 25% of the values fall below 21.82 Mbps.
* **Second Quartile (Q2)**: The second quartile, or median, is 33.7 Mbps, where 50% of the values lie below and above it.
* **Third Quartile (Q3)**: The third quartile is 61.66 Mbps, marking the median of the upper half, meaning that 75% of the values are below this threshold.
* **Mode**: The mode is not applicable here, as all values in Dataset 1 are unique.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class interval upper | Class Midpoint | Class interval lower | Count | Relative Frequency | Cumulative Relative Frequency |
| 0 | 10 | 20 | 32 | 0.203821656 | 0.203821656 |
| 20 | 30 | 40 | 61 | 0.388535032 | 0.592356688 |
| 40 | 50 | 60 | 24 | 0.152866242 | 0.74522293 |
| 60 | 70 | 80 | 12 | 0.076433121 | 0.821656051 |
| 80 | 90 | 100 | 8 | 0.050955414 | 0.872611465 |
| 100 | 110 | 120 | 8 | 0.050955414 | 0.923566879 |
| 120 | 130 | 140 | 6 | 0.038216561 | 0.961783439 |
| 140 | 150 | 160 | 2 | 0.012738854 | 0.974522293 |
| 160 | 170 | 180 | 0 | 0 | 0.974522293 |
| 180 | 190 | 200 | 2 | 0.012738854 | 0.987261146 |
| 200 | 210 | 220 | 1 | 0.006369427 | 0.993630573 |
| 220 | 230 | 240 | 0 | 0 | 0.993630573 |
| 240 | 250 | 260 | 1 | 0.006369427 | 1 |
|  |  | Total Number of Observations | 157 |  |  |

Tabular Summary of dataset - 1

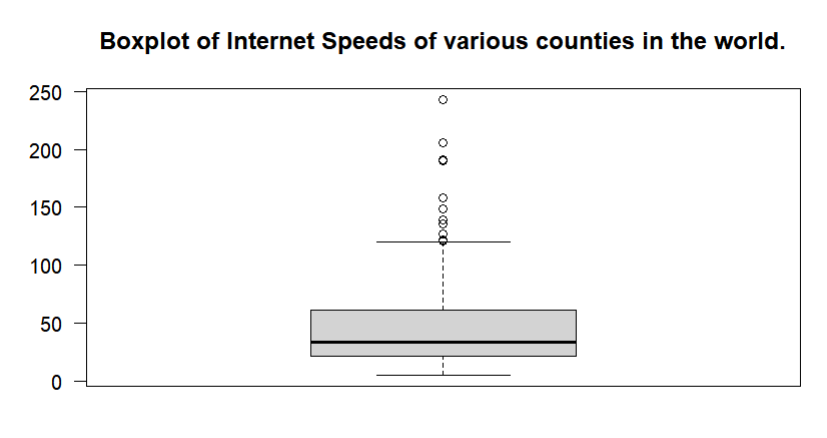
Relative Frequency Histogram for dataset -1 :

From the Relative Frequency Histogram and Tabular Summary, we can interpret that the internet speeds in Most countries , that, is 75% of countries in this dataset have an internet speed between 0 to 60 Mbps and the graph also appears to be skewed to the left.

Cumulative Relative Frequency Histogram for dataset - 1:

The cumulative relative frequency histogram shows that more countries have internet speeds below 40 Mbps than those with speeds exceeding 40 Mbps.

Box Plot for dataset - 1 :



Ogive (Line-Chart) for dataset – 1 :

Pie Chart for dataset - 1 :

Dataset - 2 : Times at which notification was pushed to Personal Android Phone Between 9 AM to 5 PM on 09/09/2024

|  |  |  |
| --- | --- | --- |
| Measurement | Time Interval (Hour in hh:mm:ss) | Time Interval (Hour in Number) |
| Sample Mean | 12:40:18 | 12.67177556 |
| Sample median | 12:46:53 | 12.78138889 |
| Max | 16:59:08 | 16.98555556 |
| Min | 09:01:36 | 9.026666667 |
| Standard Deviation | 00:24:20 | 2.249578916 |
| Sample Range | 07:57:32 | 7.958888889 |
| Quartile 1 | 10:35:56 | 10.59888889 |
| Quartile 2 | 12:46:53 | 12.78138889 |
| Quartile 3 | 14:31:37 | 14.52694444 |
| Geomean | 03:30:00 | 12.46857399 |
| Coefficient of variation | 00:00:01 | 0.177526733 |
| Mode | 09:58:54, 12:16:58 | 9.98, 12.28 |

From the measures in the table, we can interpret the following statistics:

* **Sample Mean**: The average time notifications are pushed to an Android phone is 12 hours and 40 minutes.
* **Sample Median**: The median, or central value, is also 12 hours and 40 minutes.
* **Sample Range**: The sample range is 7 hours and 57 minutes.
* **Sample Standard Deviation**: The standard deviation relative to the mean is 24 minutes.
* **First Quartile (Q1)**: The first quartile represents the median of the lower half of the dataset, indicating that 25% of the values are below 10 hours and 35 minutes.
* **Second Quartile (Q2)**: The second quartile, or median, is 12 hours and 46 minutes, with 50% of the values lying above and below this point.
* **Third Quartile (Q3)**: The third quartile is calculated to be 14 hours and 31 minutes, marking the median of the upper half and indicating that 75% of the values fall below this threshold.

Tabular Summary of dataset – 2 :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class interval upper | Class Midpoint | Class interval lower | Count | Relative Frequency | Cumulative Relative Frequency |
| 09:00:00 | 09:30:00 | 10:00:00 | 22 | 0.18 | 0.176 |
| 10:00:00 | 10:30:00 | 11:00:00 | 15 | 0.12 | 0.30 |
| 11:00:00 | 11:30:00 | 12:00:00 | 12 | 0.10 | 0.39 |
| 12:00:00 | 12:30:00 | 13:00:00 | 16 | 0.13 | 0.52 |
| 13:00:00 | 13:30:00 | 14:00:00 | 19 | 0.15 | 0.52 |
| 14:00:00 | 14:30:00 | 15:00:00 | 19 | 0.15 | 0.82 |
| 15:00:00 | 15:30:00 | 16:00:00 | 14 | 0.11 | 0.94 |
| 16:00:00 | 16:30:00 | 17:00:00 | 8 | 0.06 | 1.00 |
|  |  | Total Number of Observations | 125 |  |  |

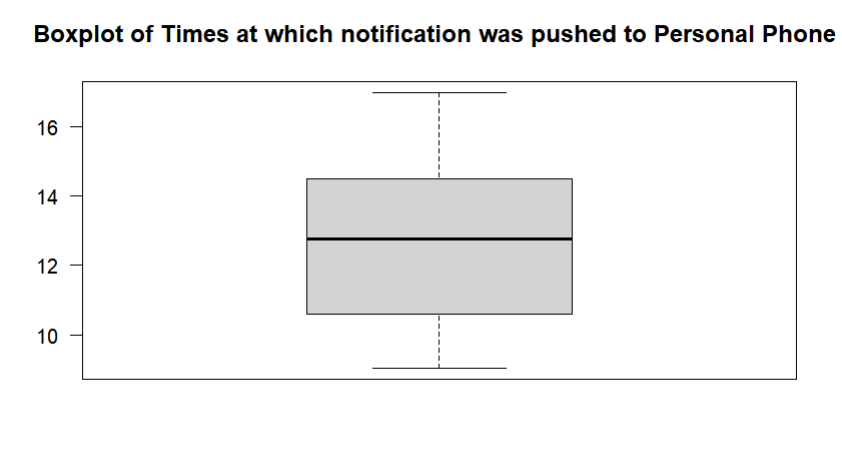
Relative Frequency Histogram for dataset – 2 :

From the relative frequency histogram and tabular summary, we can see that most notifications are pushed to Android phones between 9 and 10 AM, and there is another increase in notification frequency between 2 and 4 PM.

Cumulative Relative Frequency Histogram of dataset -2 :

The cumulative relative frequency histogram indicates that most notifications to Android phones are sent between 9 AM and 4 PM, after which the frequency of notifications decreases as the day progresses toward 5 PM.

Box Plot of dataset -2 :



Ogive (Line-Chart) of dataset -2 :

Pie Chart of dataset - 2:

**References**

Data source: <https://www.speedtest.net/global-index>

**Appendix Data set 1 - Internet Speeds of various counties in the world for the year 2022 (Measured in Mbps)**

|  |  |
| --- | --- |
| Name | Avg D Mbps |
| Cuba | 10.15 |
| Guadeloupe | 45.72 |
| Martinique | 47.84 |
| Papua New Guinea | 33.60 |
| Mauritania | 16.63 |
| Democratic Republic of the Congo | 21.41 |
| Guyana | 25.29 |
| Cameroon | 17.42 |
| Zimbabwe | 16.16 |
| Colombia | 13.60 |
| Sudan | 12.91 |
| Angola | 26.68 |
| Mexico | 28.47 |
| Cote d'Ivoire | 22.61 |
| Madagascar | 21.25 |
| Bolivia | 15.75 |
| Zambia | 15.56 |
| Peru | 22.52 |
| Dominican Republic | 27.93 |
| Trinidad and Tobago | 30.89 |
| Venezuela | 7.73 |
| Tajikistan | 13.23 |
| Afghanistan | 7.14 |
| United States Virgin Islands | 54.78 |
| Nigeria | 45.88 |
| India | 18.85 |
| Syrian Arab Republic | 17.83 |
| Haiti | 14.92 |
| Spain | 45.38 |
| United Republic of Tanzania | 15.08 |
| Iceland | 103.59 |
| Brazil | 37.21 |
| Botswana | 32.77 |
| Ghana | 11.50 |
| Yemen | 10.94 |
| Japan | 60.79 |
| Indonesia | 21.82 |
| Mozambique | 19.95 |
| Canada | 78.60 |
| Uganda | 21.61 |
| United States | 113.93 |
| Ethiopia | 30.05 |
| Iran (Islamic Republic of) | 38.71 |
| Argentina | 26.29 |
| Burma | 26.11 |
| Ecuador | 23.53 |
| Puerto Rico | 104.52 |
| United Kingdom | 71.08 |
| Italy | 53.65 |
| Albania | 37.51 |
| Somalia | 27.08 |
| Pakistan | 17.40 |
| New Zealand | 82.20 |
| Russia | 27.85 |
| Kenya | 24.73 |
| Bangladesh | 14.80 |
| Togo | 37.93 |
| Kazakhstan | 28.39 |
| French Polynesia | 25.93 |
| Paraguay | 21.21 |
| Malawi | 17.05 |
| Saudi Arabia | 121.49 |
| France | 87.24 |
| Chile | 45.14 |
| Guam | 32.83 |
| Ukraine | 27.70 |
| El Salvador | 25.30 |
| Egypt | 25.22 |
| Algeria | 19.19 |
| Sweden | 100.95 |
| Honduras | 32.98 |
| Palestine | 23.58 |
| Namibia | 18.04 |
| Portugal | 74.79 |
| Belarus | 16.61 |
| Burundi | 7.50 |
| Oman | 89.47 |
| Greece | 63.65 |
| Malaysia | 40.43 |
| Iraq | 34.16 |
| Uzbekistan | 20.24 |
| Germany | 72.66 |
| Malta | 55.61 |
| Jamaica | 38.03 |
| Romania | 37.45 |
| Georgia | 32.47 |
| Nepal | 17.96 |
| Uruguay | 41.03 |
| Turkey | 37.18 |
| Morocco | 34.10 |
| Timor-Leste | 5.45 |
| United Arab Emirates | 205.50 |
| Norway | 139.16 |
| Cyprus | 95.60 |
| Belgium | 71.90 |
| Czech Republic | 55.92 |
| Poland | 50.21 |
| Ireland | 46.87 |
| Lebanon | 30.78 |
| Lao People's Democratic Republic | 28.67 |
| Fiji | 27.99 |
| Congo | 26.50 |
| Rwanda | 17.28 |
| Libyan Arab Jamahiriya | 16.28 |
| Korea, Republic of | 242.77 |
| Croatia | 80.67 |
| South Africa | 54.32 |
| Suriname | 41.04 |
| Mongolia | 22.58 |
| Montenegro | 36.04 |
| Guatemala | 35.43 |
| Philippines | 30.84 |
| Costa Rica | 27.89 |
| Mali | 26.97 |
| Nicaragua | 24.46 |
| Sri Lanka | 18.31 |
| China | 158.09 |
| Australia | 120.63 |
| Bulgaria | 118.95 |
| Luxembourg | 100.34 |
| Austria | 77.37 |
| Bosnia and Herzegovina | 28.04 |
| Lithuania | 68.77 |
| Bhutan | 57.85 |
| Israel | 55.41 |
| Thailand | 47.46 |
| Kyrgyzstan | 25.33 |
| Panama | 21.77 |
| Cambodia | 18.23 |
| Hungary | 52.24 |
| Viet Nam | 46.18 |
| Slovakia | 44.37 |
| Republic of Moldova | 37.96 |
| Azerbaijan | 33.80 |
| Armenia | 30.56 |
| Qatar | 190.51 |
| Netherlands | 109.71 |
| Estonia | 66.61 |
| The former Yugoslav Republic of Macedonia | 62.54 |
| Senegal | 40.82 |
| Benin | 29.86 |
| Tunisia | 28.89 |
| Burkina Faso | 17.69 |
| Finland | 91.00 |
| Slovenia | 65.34 |
| Latvia | 57.59 |
| Mauritius | 34.54 |
| Denmark | 135.34 |
| Switzerland | 122.24 |
| Serbia | 45.72 |
| Singapore | 127.22 |
| Taiwan | 104.99 |
| Jordan | 21.17 |
| Bahrain | 148.51 |
| Kuwait | 190.68 |
| Brunei Darussalam | 89.05 |
| Hong Kong | 81.46 |

**Appendix Data set 1 - Times at which notification was pushed to Personal Android Phone Between 9 AM to 5 PM on 09/09/2024**

|  |
| --- |
| Times at which notification was pushed to My Phone Between 9 AM to 5 PM on 09/09/2024 |
| 14:44:12 |
| 09:30:54 |
| 15:20:16 |
| 09:14:04 |
| 12:46:53 |
| 15:37:14 |
| 10:10:17 |
| 10:41:23 |
| 13:16:57 |
| 10:27:01 |
| 09:17:39 |
| 13:03:29 |
| 11:41:43 |
| 14:25:07 |
| 12:33:19 |
| 13:21:01 |
| 14:09:24 |
| 14:37:00 |
| 14:28:07 |
| 13:00:40 |
| 15:28:48 |
| 12:43:57 |
| 14:31:49 |
| 09:01:36 |
| 09:27:57 |
| 10:04:44 |
| 12:09:14 |
| 16:49:14 |
| 15:15:09 |
| 14:04:44 |
| 10:23:17 |
| 10:35:56 |
| 09:11:03 |
| 09:09:51 |
| 14:43:29 |
| 15:21:03 |
| 13:47:20 |
| 14:13:18 |
| 11:59:36 |
| 16:59:08 |
| 09:22:46 |
| 09:43:01 |
| 11:21:46 |
| 11:47:54 |
| 10:00:16 |
| 09:21:15 |
| 12:26:08 |
| 10:12:15 |
| 12:21:53 |
| 15:07:33 |
| 11:20:46 |
| 14:53:54 |
| 09:58:54 |
| 14:22:48 |
| 12:31:55 |
| 09:42:11 |
| 09:44:24 |
| 13:28:43 |
| 09:22:16 |
| 16:13:56 |
| 11:48:40 |
| 14:52:24 |
| 10:52:53 |
| 11:48:04 |
| 10:42:00 |
| 15:23:20 |
| 16:04:49 |
| 14:31:54 |
| 12:16:58 |
| 13:34:22 |
| 12:19:13 |
| 16:57:52 |
| 09:33:04 |
| 13:00:42 |
| 14:58:03 |
| 15:27:22 |
| 13:33:52 |
| 10:18:53 |
| 09:03:39 |
| 12:06:45 |
| 11:11:52 |
| 16:02:03 |
| 14:00:21 |
| 13:37:51 |
| 14:12:15 |
| 13:39:39 |
| 09:48:20 |
| 11:55:44 |
| 09:07:56 |
| 11:37:34 |
| 10:49:54 |
| 12:54:17 |
| 13:54:08 |
| 15:11:06 |
| 09:30:45 |
| 15:41:52 |
| 12:43:28 |
| 13:33:17 |
| 13:42:38 |
| 09:58:54 |
| 12:49:02 |
| 12:16:58 |
| 14:03:41 |
| 16:01:53 |
| 13:50:26 |
| 11:15:09 |
| 13:15:18 |
| 14:55:44 |
| 12:00:05 |
| 13:20:52 |
| 13:23:25 |
| 14:31:37 |
| 12:39:09 |
| 15:50:54 |
| 09:05:46 |
| 10:55:34 |
| 10:34:07 |
| 15:54:38 |
| 09:12:18 |
| 11:57:32 |
| 15:09:41 |
| 15:58:08 |
| 16:27:18 |
| 13:48:26 |
| 10:19:28 |