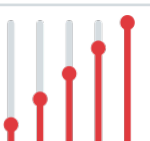
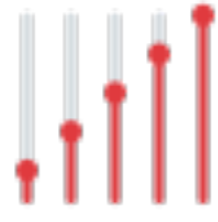


Measuring Broadband America

Data Processing Flow, Ninth Report

The following diagram depicts how raw data was processed for the production of the Measuring Broadband America Ninth Report





Data Processing Flow Steps:

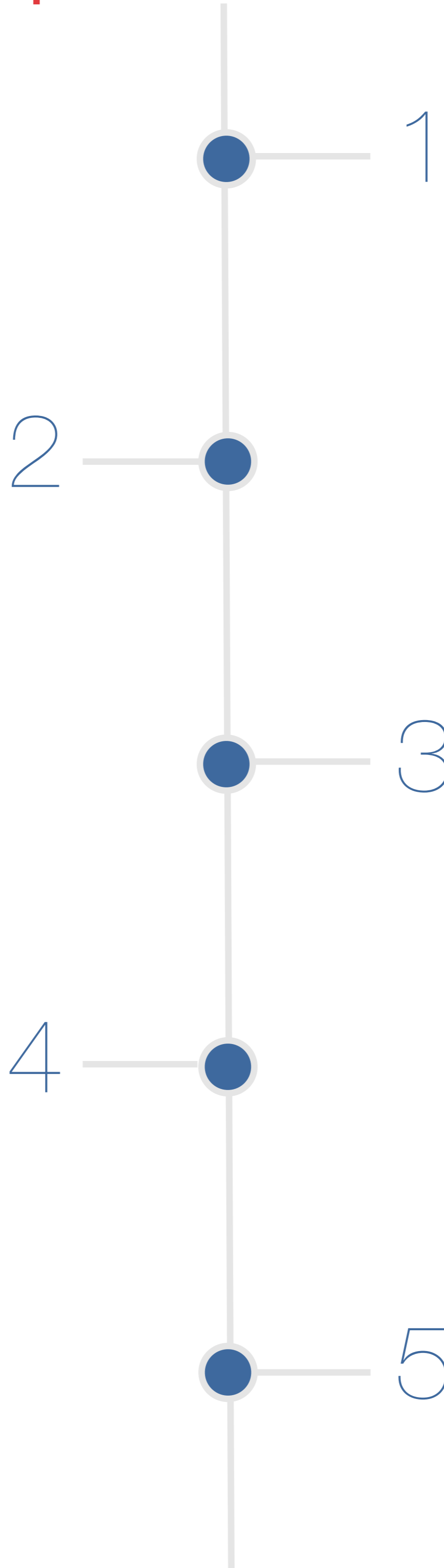
Clean Data

Data is cleaned. This includes removing measurements when a user changed ISP/product during the period.

Anomalies and significant outliers are also removed at this point. SQL scripts used to clean the data are published under the name 'SQL Cleanup Scripts'.

SPSS Processing

The per-device, per-metric CSV data is imported into an Excel workbook for further processing. It is at this point that the data is coupled with the unit profile data. ISPs and products with low sample sizes are excluded at this point.



1

Raw Data

Raw data for the chosen period is collected from the measurement database. The ISPs and products that panellists were on are exported to a 'unit profile' file. ISP and product assignments are obtained through the validation process operated in conjunction with ISPs.

3

Per-device Aggregates

Per-device aggregates are generated for each metric. It is at this point that the 7-11pm averages are computed, along with trimmed means, medians and various percentiles for each metric. The SQL scripts used are published under the name 'SQL Processing Scripts'.

4

5

Excel Tables / Charts

Summary data tables and charts in Excel are produced using the same processing workbook. This workbook is published under the name 'Processing Workbook'. This directly produces all of the charts and statistical averages files.

