Internet Data Portal

Jay Daley, .nz
https://idp.nz

The Internet Data Portal is a collection of open data about the Internet in NZ. Using this site, you can download, explore or comment on the data. We hope you find this site useful and welcome any feedback you have.

Open Data

- Domain Names
- Internet Connectivity
- Global Rankings
- Users and Use
Three years of data

- Repository for Internet data relating to NZ
- All open data
- Category 1 – domain names
  All our .nz data we make public
- Category 2 – Internet connectivity
  ISPs and Connections in NZ
- Category 3 – Global rankings
  Any internet related data that ranks countries
- Category 4 - Users and use
  Primarily surveys of internet users
Current .nz datasets

- .nz Zone Scan
  - Run monthly
  - Looks for 50+ indicators
  - Aggregate stats uploaded after each scan
- .nz Registration data
  - Transaction per day by type of transaction
  - Domains in each state
  - Aggregate stats uploaded each day
- SSL/TLS Certificates in .nz
- .nz nameserver packet capture analysis
TTLs from zone scan
Underlying data

<table>
<thead>
<tr>
<th>Classification</th>
<th>Metric</th>
<th>Count Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTL MX</td>
<td>[0, 5m[</td>
<td>Domains</td>
<td>5,276</td>
</tr>
<tr>
<td>TTL MX</td>
<td>5m</td>
<td>Domains</td>
<td>35,362</td>
</tr>
<tr>
<td>TTL MX</td>
<td>]5m, 10m[</td>
<td>Domains</td>
<td>447</td>
</tr>
<tr>
<td>TTL MX</td>
<td>10m</td>
<td>Domains</td>
<td>56,298</td>
</tr>
<tr>
<td>TTL MX</td>
<td>]10m, 15m[</td>
<td>Domains</td>
<td>171</td>
</tr>
<tr>
<td>TTL MX</td>
<td>15m</td>
<td>Domains</td>
<td>3,449</td>
</tr>
<tr>
<td>TTL MX</td>
<td>]15m, 1h[</td>
<td>Domains</td>
<td>5,497</td>
</tr>
<tr>
<td>TTL MX</td>
<td>1h</td>
<td>Domains</td>
<td>131,345</td>
</tr>
<tr>
<td>TTL MX</td>
<td>]1h, 4h[</td>
<td>Domains</td>
<td>6,028</td>
</tr>
<tr>
<td>TTL MX</td>
<td>4h</td>
<td>Domains</td>
<td>68,775</td>
</tr>
</tbody>
</table>
Lessons learned

• **Product**
  API access – upload and download
  Reliability and speed for regular uploads
  Catalogue / persistent URLs / data formats

• **Long vs Wide**
  Long is much better for data analysis
  Requires data pre-processing and discipline

• **Changing source data**
  Employ student regularly to do messy stuff

• **Eat our own dog food**
  Remember – all new datasets go on portal
  Need data internally - use the portal
Thanks

Find it at https://idp.nz

Contact: Jay Daley, jay@nzrs.net.nz
www.nzrs.net.nz