

Internet Data Portal

Jay Daley, .nz

https://idp.nz

The screenshot displays the Internet Data Portal website. At the top left is the NZRS logo and the title "Internet Data Portal". On the top right are navigation links for "Home" and "Catalogue", a search icon, and a "Sign In" button. Below the header is a descriptive paragraph: "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." To the right of this text is a search input field with the placeholder "Search". Below the text are three preview images: a bar chart titled "About this portal" with a legend for categories like "Very Small (1-100)", "Small (101-1000)", "Medium (1,001-10,000)", and "Large (10,001-100,000)"; a "Search & Browse Datasets and Catalogue" section listing "Internet Service Providers - The percentages of Internet" and "Internet Connections: Data - The number of internet conn"; and a "Using this site" section with a "Conditions" sub-header. At the bottom of the screenshot is a blue banner with the text "Open Data" and four icons representing "Domain Names", "Internet Connectivity", "Global Rankings", and "Users and Use".



Three years of data

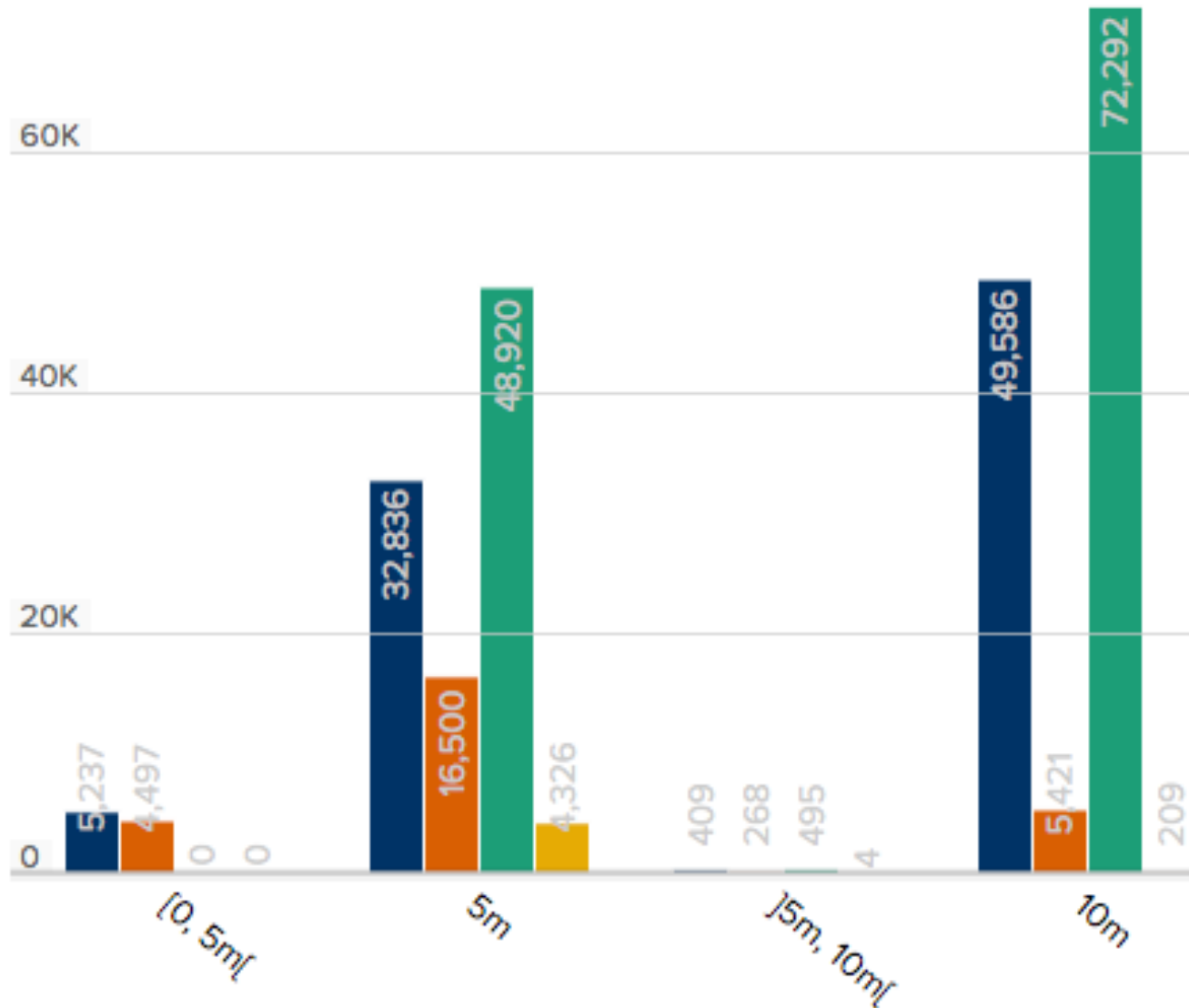
- Repository for Internet data relating to NZ
- All open data
- Category 1 – domain names
All our .nz data we make public
- Category 2 – Internet connectivity
ISPs and Connections in NZ
- Category 3 – Global rankings
Any internet related data that ranks countries
- Category 4 – Users and use
Primarily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

Classification	Metric  	Count Type	Count  
TTL MX	[0, 5m[Domains	5,276
TTL MX	5m	Domains	35,362
TTL MX]5m, 10m[Domains	447
TTL MX	10m	Domains	56,298
TTL MX]10m, 15m[Domains	171
TTL MX	15m	Domains	3,449
TTL MX]15m, 1h[Domains	5,497
TTL MX	1h	Domains	131,345
TTL MX]1h, 4h[Domains	6,028
TTL MX	4h	Domains	68,775

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](mailto:jay@nzrs.net.nz)
www.nzrs.net.nz