Public Disclosure Request Type Analysis

By Angela Gonzalez-Curci for WSDOT as part of University of Washington/Open Data Literacy Internship

# Goals:

Find most frequently requested Washington State Department of Transportation records.

Assess the applicability of the classification scheme developed by Kenneth Lee’s Public Disclosure Project Final Report to unstructured request data.

# Analysis Strategy:

Take sample of full dataset and code the requestor responses recorded in the Request Item field of the PRD dataset. Using Kenneth Lee’s classification scheme, with additional secondary and tertiary record types added to the classification to fill gaps between the scheme and the responses using literary warrant, derived from the text of request itself.

After classification, frequency of classes and subclasses were taken with Excel Countif function. The percentage of each class and subclass was then calculated by dividing class count by the total sample count.

# Classification Scheme:

In 2014, Kent State Graduate Student Kenneth Lee proposed an extension to already existing record classification schedules based on a big bucket strategy used multiple federal agencies. Big bucket strategy attempts to simplify record classification schemes by first classifying records by large overview categories which are created by and easily understood by outside users. Using Washington State Department of Transportation own record classification schedules. Ken’s schedule is attached as an index to this report.

# Method

Data analysis was done in an Excel spreadsheet titled PDR Request Type Analysis. A sample of 1040 entries from the PDR dataset.

After classification, frequency of classes and subclasses were taken with Excel Countif function. The percentage of each class and subclass was then calculated by dividing class count by the total sample count.

Since the sample is such a large portion of the whole population of the dataset, the margin of error needed to be adjusted with the Finite Population Corrector:

Finite Population Corrector =

N = Population Size

n = Sample Size

With this adjustment the Margin of Error was calculated for each class and subclass:

Margin of Error = FPC \* Z-Value \*

FPC = Finite Population Corrector

Z-Value for 99% confidence level = 2.58

P = Percentage of Sample

n= Sample Size

# Analysis:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classification | Count | | Percentage of Sample | "+/-" |
| Administration | | **34** | **3.82%** | **0.05%** |
|  | |  | 0.00% | 0.00% |
| Communication \* | | **14** | **1.57%** | **0.03%** |
| External \* | | 4 | 0.45% | 0.02% |
| Internal \* | | 5 | 0.56% | 0.02% |
| Personal \* | | 5 | 0.56% | 0.02% |
|  | |  |  |  |
| Legal \* | | **14** | **1.57%** | **0.03%** |
| Public Record Request \* | | 9 | 1.01% | 0.03% |
| Case \* | | 5 | 0.56% | 0.02% |
|  | |  |  |  |
| Structure \* | | **3** | **0.34%** | **0.02%** |
| Organization Tables \* | | 2 | 0.22% | 0.01% |
| DOT Shuttle Express Office Locations \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Policy \* | | **3** | **0.34%** | **0.02%** |
| Executive Order \* | | 2 | 0.22% | 0.01% |
| Best Practice \* | | 1 | 0.11% | 0.01% |
|  | |  | 0.00% | 0.00% |
| Finance | | **18** | **2.02%** | **0.04%** |
|  | |  |  |  |
| Accounting | | **11** | **1.23%** | **0.03%** |
| Payroll | | 9 | 1.01% | 0.03% |
| Invoices | | 2 | 0.22% | 0.01% |
| Receiving | | 0 | 0.00% | 0.00% |
|  | |  |  |  |
| Budget | | **7** | 0.79% | 0.02% |
| Report | | 6 | 0.67% | 0.02% |
| Budget Plan \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Human Resources | | **5** | **0.56%** | 0.02% |
|  | |  | 0.00% | 0.00% |
| Application Information | | 5 | 0.56% | 0.02% |
| Investigation Notes \* | | 0 | 0.00% | 0.00% |
| Personnel Files | | 0 | 0.00% | 0.00% |
|  | |  |  |  |
| Information Technology | | **3** | **0.34%** | 0.02% |
|  | |  |  |  |
| Information Technology | | **2** | **0.22%** | 0.01% |
| Ops Log \* | | 1 | 0.11% | 0.01% |
| Dashboard Information \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Data | | **1** | **0.11%** | 0.01% |
| GIS \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Project Management | | **1** | **0.11%** | 0.01% |
|  | |  |  |  |
| Project Management | | **1** | **0.11%** | 0.01% |
| Independent Project Manager's Technical Report \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Transportation Operations | | **730** | **81.93%** | **0.11%** |
|  | |  |  |  |
| Bridge | | **14** | **1.57%** | **0.03%** |
| Plans \* | | 11 | 1.23% | 0.03% |
| Inspection \* | | 2 | 0.22% | 0.01% |
| List \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Collision | | **310** | **34.79%** | 0.13% |
| *Data* \* | | **284** | **31.87%** | 0.13% |
| Collision | | 69 | 7.74% | 0.08% |
| Crash | | 214 | 24.02% | 0.12% |
| Incident | | 1 | 0.11% | 0.01% |
| *Report* \* | | **13** | **1.46%** | 0.03% |
| Accident Report | | 5 | 0.56% | 0.02% |
| Collision Report | | 2 | 0.22% | 0.01% |
| Report | | 6 | 0.67% | 0.02% |
|  | |  |  |  |
| Construction | | **199** | **22.33%** | 0.12% |
| All \* | | 22 | 2.47% | 0.04% |
| As-built | | 10 | 1.12% | 0.03% |
| Channelization Plans \* | | 1 | 0.11% | 0.01% |
| Drainage Plans \* | | 1 | 0.11% | 0.01% |
| Intersection Plan \* | | 2 | 0.22% | 0.01% |
| Plan Set | | 31 | 3.48% | 0.05% |
| Right-of-Way | | 105 | 11.78% | 0.09% |
| Slide Remediation Plans \* | | 0 | 0.00% | 0.00% |
| Easement \* | | 2 | 0.22% | 0.01% |
| Design | | 2 | 0.22% | 0.01% |
| Design Documentation Package | | 1 | 0.11% | 0.01% |
| Design Standards \* | | 1 | 0.11% | 0.01% |
| Drainage Report | | 1 | 0.11% | 0.01% |
| Post-Construction Building Survey Report \* | | 1 | 0.11% | 0.01% |
| Engineering Drawings \* | | 1 | 0.11% | 0.01% |
| Field Note Records \* | | 2 | 0.22% | 0.01% |
| Highway Data\* | | 1 | 0.11% | 0.01% |
| Highway Maps\* | | 1 | 0.11% | 0.01% |
| Map\* | | 6 | 0.67% | 0.02% |
| Latest Value Engineering Study\* | | 0 | 0.00% | 0.00% |
| Permit\* | | 2 | 0.22% | 0.01% |
| Plane Coordinates\* | | 1 | 0.11% | 0.01% |
| Professional Engineer Certification of Review\* | | 2 | 0.22% | 0.01% |
| Road Survey\* | | 1 | 0.11% | 0.01% |
| Sound Studies\* | | 1 | 0.11% | 0.01% |
| Turnback\* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Contract\* | | **42** | **4.71%** | 0.06% |
| Ad and Award | | 20 | 2.24% | 0.04% |
| Bid | | 20 | 2.24% | 0.04% |
| Change Order | | 1 | 0.11% | 0.01% |
| Estimates | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Ferries | | **46** | **5.16%** | 0.06% |
| Fare Data\* | | 3 | 0.34% | 0.02% |
| Incident Report | | 38 | 4.26% | 0.06% |
| Performance Data \* | | 3 | 0.34% | 0.02% |
| Public Comments \* | | 1 | 0.11% | 0.01% |
| Video Record \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Maintenance | | **21** | **2.36%** | 0.04% |
| Highway Pass Information | | 20 | 2.24% | 0.04% |
| Operations | | 1 | 0.11% | 0.01% |
|  | |  | 0.00% | 0.00% |
| Materials Lab | | **13** | **1.46%** | 0.03% |
| Toxicity Characteristic Leaching Procedure\* | | 1 | 0.11% | 0.01% |
| Boring Log | | 8 | 0.90% | 0.03% |
| Geotechnical | | 3 | 0.34% | 0.02% |
| Roadway Testing | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Materials Management | | **1** | 0.11% | 0.01% |
| Purchasing | | 1 | 0.11% | 0.01% |
| Environment | | 0 | 0.00% | 0.00% |
|  | |  | 0.00% | 0.00% |
| Rail | | **5** | **0.56%** | 0.02% |
| Light Rail \* | | 1 | 0.11% | 0.01% |
| Amtrak or Other Rail Related Information | | 4 | 0.45% | 0.02% |
|  | |  |  |  |
| Tolling | | **4** | **0.45%** | 0.02% |
| Information | | 2 | 0.22% | 0.01% |
| Customer Service Center | | 2 | 0.22% | 0.01% |
|  | |  |  |  |
| Traffic | | **69** | **7.74%** | 0.08% |
| Speed Traffic Survey\* | | 4 | 0.45% | 0.02% |
| Traffic Control Plan\* | | 1 | 0.11% | 0.01% |
| Traffic Data \* | | 19 | 2.13% | 0.04% |
| Traffic Camera | | 26 | 2.92% | 0.05% |
| Speed Limit Log\* | | 6 | 0.67% | 0.02% |
| Signal Plan \* | | 5 | 0.56% | 0.02% |
| Electronic Sign Log\* | | 3 | 0.34% | 0.02% |
| Signal Log\* | | 2 | 0.22% | 0.01% |
| Traffic Projection Study\* | | 2 | 0.22% | 0.01% |
| Traffic Study\* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Weigh Station | | **6** | 0.67% | 0.02% |
| Haul Permits \* | | 1 | 0.11% | 0.01% |
| License Plate Readers \* | | 180 | 16.81% | 0.09% |
| List of Transponders \* | | 4 | 0.45% | 0.02% |
| Location Data \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| Transportation Asset Management | | **15** | 1.68% | 0.04% |
|  | |  |  |  |
| Facilities | | **0** | 0.00% | 0.00% |
| Commercial Vehicles | | 0 | 0.00% | 0.00% |
|  | |  |  |  |
| Real Estate | | **15** | 1.68% | 0.04% |
| Deed | | 10 | 1.12% | 0.03% |
| Appraisal | | 1 | 0.11% | 0.01% |
| Lease \* | | 3 | 0.34% | 0.02% |
| SWR Real Estate Fact Finding Report \* | | 1 | 0.11% | 0.01% |
|  | |  |  |  |
| N/A | | 13 | 1.46% | 0.03% |

\* Added Class

# Results

While request for internal information regarding the workings of WSDOT may take be the most complicated and time consuming requests to deliver for Coordinators, the vast majority of requests are for records concerning transportation.

The dominance of Transportation Operations as a primary record type also speaks to one the flaws in applying the proposed classification scheme to the record requests. There is a major difference in the types of records WSDOT creates and maintains and those requested by the public. A better schema would take this difference into account.

|  |  |  |  |
| --- | --- | --- | --- |
| Primary Record Type | Count | % | Margin of Error |
| Administration | 34 | 3.11% | 0.04% |
| Finance | 19 | 1.74% | 0.03% |
| Human Resources | 19 | 1.74% | 0.03% |
| Information Technology | 3 | 0.27% | 0.01% |
| Project Management | 1 | 0.09% | 0.01% |
| Transportation Operations | 914 | 83.55% | 0.08% |
| Transportation Asset Management | 19 | 1.74% | 0.03% |

Three record types make up about 45% of all records requests: Crash Data, Right of Way, and License Plate Readers. This confirms much of the anecdotal experience of Public Disclosure Coordinators. All three of these types were mentioned by coordinators in my preliminary interviews with them.

|  |  |  |  |
| --- | --- | --- | --- |
| Top Record Types | Count | % of Sample | Margin of Error |
| Crash Data | 214 | 19.56% | 0.09% |
| Right-of-Way | 105 | 9.60% | 0.07% |
| License Plate Readers | 180 | 16.45% | 0.08% |
| Total | **499** | **45.61%** |  |
|  |  |  |  |

Records held by Crash Data and Requests for the data and reports managed by the WSDOT Crash Data and Reporting Branch made up about 28% of the total.

|  |  |  |  |
| --- | --- | --- | --- |
| Record held by Crash Data | Count | % of Sample | Margin of Error |
| Data | **284** | **25.96%** | **0.10%** |
| Collision | 69 | 6.31% | 0.06% |
| Crash | 214 | 19.56% | 0.09% |
| Incident | 1 | 0.09% | 0.01% |
| Report | **13** | **1.19%** | **0.02%** |
| Accident Report | 5 | 0.46% | 0.02% |
| Collision Report | 2 | 0.18% | 0.01% |
| Report | 6 | 0.55% | 0.02% |
| Total | **310** | **28.34%** | **0.10%** |

# Big Bucket Classification Schedule:

