On the first level (main headings), only Business Economics and Finance stands out as having a high relative occurrence count. Its branch represents over 50% of the total UNTL-BS occurrences on the Portal, yet contains only 8% of the narrower terms. When taken down to the second level (main heading and one narrower term), this is shown to be the result of three major outliers in this branch: Advertising, Communications, and Journalism. Advertising and Journalism each comprise about 28% of the occurrences on the BEF branch, yet neither contains any narrower terms. Communications, on the other hand, comprises 37%. Among its three narrower terms, two are very well-used and two are not. Newspapers comprises over 77% of the total occurrences under Communications, and Media comprises 22%.

Next, we need to open up all of the second level entries, because the extreme skew of the BEF branch prevents any other outliers from becoming evident at the top level.

Agriculture

* The main heading alone accounts for over 6% of the occurrences on the whole branch. This is relatively high for a main heading.
* Domestic Animals has a particularly high relative occurrence count, comprising almost 44% of the total occurrences on the branch. It does have 10 narrower terms in its branch, but relative to its total occurrence count this doesn’t seem to be enough. Drilling one level further down, it is revealed that the bulk of the occurrences on this branch are split three ways.
  + First, the main heading and one narrower term is responsible for over 34% of the total. This is strong evidence for the need for more narrower terms at this level – if it is not revealed that many of these occurrences are erroneous.
  + Second, Cattle accounts 25% of all occurrences yet has no narrower terms. This would seem to call for further investigation.
  + Third, Horses accounts for almost 28% of all occurrences. It has one narrower term (Tack), but looking at the straight UNTL-BS counts reveals that this narrower term only siphons away 489 occurrences away from the broader term, which comprises 4,364 by comparison.

Architecture

* Buildings is a pretty obvious outlier, comprising almost 74% of the total occurrences on the branch. Furthermore, digging down one level reveals that it does not have any narrower terms, making it a prime target for further investigation.

Arts and Crafts

* Music is the most obvious outlier, comprising over half of the occurrences on the branch. It does have five narrower terms under it, but drilling down to the second level it is obvious that they do not do much to siphon off the occurrences evenly: the main heading and single narrower term alone are responsible for 84%. Unless many of these occurrences are the result of an error or mis-application, this would strongly suggest that Arts and Crafts – Music should have additional narrower terms.
* Theatre is the second-most obvious, comprising 19% of the total occurrences yet having only one narrower term. This narrower term (Motion Pictures) does hold its own, siphoning off 24% of the occurrences in the branch. But with the remaining 75% still clinging to the main heading + single narrower term, it may be that Arts and Crafts – Theatre could use additional narrower terms.
* Dance and the main heading on its own both show a bit of a delta value, but not high enough to demand further investigation.

Education

* Schools is the only term to really stand out at the second level. When expanded, it becomes clear that this term is heavily weighted toward one of its narrower terms, Students.

Immigration

* Both the main heading and its single narrower term display as having high deltas in the spreadsheet, but this is merely a side effect of the small size of the branch. With only two terms, each comprises 50% of the total and has no narrower terms.

Landscape and Nature

* Geography and Maps is really the only term with a high relative occurrence count – it comprises 62% while having no narrower terms. This is a strong indication that it merits further investigation.

Literature

* The main heading on its own comprises over 62% of the occurrences, which might suggest that it needs additional narrower terms. Except that, as with Immigration, this delta value is skewed by the small size of the branch. 1,641 total occurrences is not enough to warrant additional terms.

Military and War

* Personnel is clearly the heaviest term in this branch, comprising 20% of the occurrence count while having only one narrower term. This narrower term (Veterans) does siphon off 22% of the occurrences on the branch, but the 78% that remain on the level above suggest that this might be a term to investigate further.

People

* Within this branch, only Individuals has a high relative occurrence count, despite its 16 narrower terms. When drilling down to the next level, it is obvious that these narrower terms do very little: each comprises less than 0.2% of the total occurrences, whereas the main heading and one narrower term account for over 99.5%. This is absolutely a term to examine further, unless it can be determined that its abnormally high concentration of occurrences is due to a mis-application.

Places

* This branch has no apparent skew – “United States” is a very obvious outlier, but since it is so in both occurrence count and in narrower terms, it balances itself out

Religion

* No real skew here either

Science and Technology

* It is very clearly the main heading that holds the bulk of the skew in this branch, accounting for over 66% of the total occurrences.

Social Life and Customs

* Correspondence has a high relative occurrence rate (16% of the total occurrences on the branch with no narrower terms). This becomes even more pronounced when we observe that two other narrower terms on this branch that have similar occurrence rates (Clubs and Organizations, and Customs) have 20 narrower terms each.

Sports and Recreation

* Football is the only term that appears to have a slightly high relative occurrence rate (21% with no narrower terms).

Opening up all of the lower levels revealed a few additional terms of interest on branches that appeared to be balanced on higher levels:

* People – Ethnic Groups – African Americans has a significantly high relative occurrence count (51% with no narrower terms), yet was effectively masked by the counterbalance of the term American Indians, which has a high number of narrower terms (26) with very few occurrences.
* People – Groups: its single narrower term causes the branch to appear balanced when viewed at a higher level, but digging down reveals that this narrower term siphons off only 1% of the occurrences. With the remaining 99% on the broader term, and over 22,000 occurrences in total, this term warrants further examination.
* Social Life and Customs – Clubs and Organizations: its many narrower terms don’t siphon off nearly enough of the occurrences, leaving over 73% still on the broader term.

Taken together, this exercise suggests that the following 20 terms be examined more carefully:

1. Agriculture – Domestic Animals
2. Agriculture – Domestic Animals – Cattle
3. Agriculture – Domestic Animals – Horses
4. Architecture – Buildings
5. Arts and Crafts – Music
6. Arts and Crafts – Theatre
7. Business, Economics and Finance - Advertising
8. Business, Economics and Finance - Communications – Media
9. Business, Economics and Finance - Communications – Newspapers
10. Business, Economics and Finance – Journalism
11. Education – Schools – Students
12. Landscape and Nature – Geography and Maps
13. Military and War – Personnel
14. People – Ethnic Groups – African Americans
15. People – Groups
16. People – Individuals
17. Science and Technology
18. Social Life and Customs – Clubs and Organizations
19. Social Life and Customs – Correspondence
20. Sports and Recreation – Football

Third-level terms (main heading + two narrower terms) with high relative occurrence rates, which don’t warrant inclusion on the terms of interest list because their absolute occurrence counts are too low:

* Agriculture – Farm Equipment (57%, but only 2,067)
* Architecture – Civil Works – Streets and Roads (73%, but only 10,622)
* Architecture – Landmarks – Historic Markers (58%, but only 2,791)
* Arts and Crafts – Paintings (88%, but only 2,101)
* Business, Economics and Finance – Factories (91%, but only 1,099)
* Business, Economics and Finance – Service Industries – Photography (83%, but only 8,077)
* Business, Economics and Finance – Stores (45%, but only 5,327)
* Education – Events (79%, but only 6,835)
* Government and Law – Civil Servants – Firefighters (72%, but only 8,385)
* Government and Law – Legal Documents (37%, but only 2,748)
* Government and Law – Legal Documents – Probate (43%, but only 3,189)
* Government and Law – Money (80%, but only 12,796)
* Government and Law – Politics (64%, but only 10,282)
* Government and Law – Vital Records – Deaths (90%, but only 9,905)
* Landscape and Nature – Natural Disasters – Fires (56%, but only 6,750)
* Military and War – Forts – Fort Hood (80%, but only 2,257)
* Military and War – United States Armed Forces – United States Army (53%, but only 9,464)
* Military and War – Wars – World War II (59%, but only 16,337)
* Military and War – Weapons (90%, but only 1,903)
* Places – Canada (62%, but only 296)
* Places – France (69%, but only 56)
* Places – Spain – Texas (96%, but only 625)
* Religion – Churches (74%, but only 8,141)
* Religion – Missions (57%, but only 276)
* Science and Technology – Tools (99%, but only 3,877)
* Social Life and Customs – Clothing (65%, but only 7,702)
* Social Life and Customs – Competitions (69%, but only 3,808)
* Social Life and Customs – Food and Cooking (67%, but only 5,261)
* Social Life and Customs – Furnishings – Furniture (46%, but only 2,064)
* Social Life and Customs – Hobbies – Collections (62%, but only 440)
* Social Life and Customs – Pets – Dogs (73%, but only 2,292)
* Social Life and Customs – Popular Culture – Television (64%, but only 3,692)
* Sports and Recreation – Games (99%, but only 1,069)
* Sports and Recreation – Riding – Horses (96%, but only 8,030)
* Government and Law – Law Enforcement – Police has 46% of the occurrences under Law Enforcement, but the fact that it shares its branch with other terms that comprise 29% (Investigations) and 13% (Jails and Prisons) means that the occurrence count is actually relatively well-spread-out.
* Likewise, Military and War – Transportation – Aviation and Military and War – Transportation – Ships might both look like they have high relative occurrence rates within their branch (46% and 41%, respectively), but the fact that they divide the occurrences up between them means that the branch is actually pretty well-balanced.

There are also several terms that *technically* have a delta score of 1 – but only because they are the only narrower term on their branch. Of these, only Architecture – Buildings has a high enough absolute occurrence count (115,444) to warrant inclusion.