

Appendix I

This appendix contains a table of key classes and properties of the Canadian Census Ontology (the OWL file can be found here at our GitHub page: <https://github.com/EnterpriseIntegrationLab/CKGN/tree/main/UniversityOfToronto/Census/Ontologies>)

Class	Property	Value restriction
CensusProfile	hasCharacteristic	only Characteristic
	hasTime	only time:DateTimeInterval
Characteristic	rdfs:subClassOf	21972:Indicator
	hasLocation	5087:CityOntologyThing
	hasPart	only Characteristic
	cardinality_of	only 21972:Population
	value	only 21972:Measure
CensusTract	subClassOf	5087:CityAdministrativeArea
	5087:properPartOf	only 5087:CityAdministrativeArea
	5087:hasProperPart	only 5087:CityAdministrativeArea
	5087:hasLocation	only 5087:Location

Appendix II Additional SPARQL query examples

This section contains the remainder of the competency questions' translations into SPARQL and the results returned by the knowledge graph.

CQ1: Which NIAs have the highest number of unemployed individuals?

The following query finds the value of the characteristic "Unemployed" for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood (sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood toronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?unemployed a uoft:Unemployed2016;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
  ?population a ?populationclass.
  ?populationclass iso21972:defined_by uoft:PersonUnemployed2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 1) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of unemployed residents (2665) while Neighbourhood 139 (Scarborough Village) has the lowest number of unemployed residents (980).

CQ2: What industries do NIA residents work in?

Supplementary Figure 1: Results of the unemployed query

	neighborhood	sumvalue
1	#neighborhood137	"2665" ^{xs:integer}
2	#neighborhood136	"1515" ^{xs:integer}
3	#neighborhood135	"995" ^{xs:integer}
4	#neighborhood139	"980" ^{xs:integer}

Supplementary Figure 2: Results of the industry query

	neighborhood	sumvalue
1	.#neighborhood137	"2800""xsd:integer
2	.#neighborhood136	"1025""xsd:integer
3	.#neighborhood135	"895""xsd:integer
4	.#neighborhood139	"620""xsd:integer

- For example, a NIA where many residents work in the manufacturing sector might benefit from IPC training and certification.

The following query finds the value of the characteristic “31-33 Manufacturing” (which represents the number of people who work in the manufacturing sector according to NAICS codes 31-33) for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood toronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?manufacturing a uoft:31-33Manufacturing2016;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
  ?population a ?populationclass.
  ?populationclass iso21972:defined_by uoft:Person31-33Manufacturing2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 2) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents (2800) who work in the manufacturing industry while Neighbourhood 139 (Scarborough Village) has the lowest number of residents (620) who work in the manufacturing industry.

CQ4: Which NIAs have the most residents with a high shelter-cost-to-income ratio (30+ %)?

The following query finds the value of the characteristic “Spending 30 Percent or More of Income on Shelter Costs” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood toronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?highsheltercost a uoft:Spending30PercentOrMoreOfIncomeOnShelterCosts2016;
    uoft:hasLocation ?censustract;
    iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
}
GROUP BY ?neighbourhood

```



Supplementary Figure 3: Results of the high shelter cost query

	neighborhood	sumvalue
1	.#neighborhood137	*6605**xsd:integer
2	.#neighborhood136	*3230**xsd:integer
3	.#neighborhood139	*2460**xsd:integer
4	.#neighborhood135	*2090**xsd:integer

Supplementary Figure 4: Results of the no degree/diploma query

	neighborhood	sumvalue
1	.#neighborhood137	*8850**xsd:integer
2	.#neighborhood136	*4990**xsd:integer
3	.#neighborhood139	*2955**xsd:integer
4	.#neighborhood135	*2465**xsd:integer

ORDER BY DESC(?sumvalue)

By running the SPARQL query, we can see (Supplementary Figure 3) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents with a high shelter-cost-to-income ratio (6605) while Neighbourhood 135 (Morningside) has the lowest number of residents with a high shelter-cost-to-income ratio (2090).

CQ5: Which NIAs have the most residents who do not have a postsecondary degree nor high school diploma?

The following query finds the value of the characteristic “No Certificate Diploma or Degree” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood (sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137,
        toronto:neighbourhood139 ) )
    ?nodegree a uoft:NoCertificateDiplomaOrDegree2016;
        uoft:hasLocation ?censustract;
        iso21972:cardinality_of ?population;
        iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by
        uoft:PersonNoCertificateDiplomaOrDegree2016.
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 4) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents who do not have a postsecondary degree (8850) while Neighbourhood 135 (Morningside) has the lowest number of residents who do not have a postsecondary degree (2465).

CQ6: Which NIAs have the most residents who know neither English nor French?

The following query finds the value of the characteristic “Neither English Nor French” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>

```

Supplementary Figure 5: Results of the neither English nor French query

	neighborhood	sumvalue
1	.#neighborhood137	"2575"^^xsd:integer
2	.#neighborhood136	"600"^^xsd:integer
3	.#neighborhood139	"525"^^xsd:integer
4	.#neighborhood135	"485"^^xsd:integer

```

SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137,
        toronto:neighbourhood139 ) )
    ?noenglishfrench a uoft:NeitherEnglishNorFrench2016;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by
    uoft:PersonNeitherEnglishNorFrench2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 5) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents who know neither English nor French (2575) while Neighbourhood 135 (Morningside) has the lowest number of residents who know neither English nor French (485).

CQ7: What are the largest visible minority groups in NIAs?

The following query finds the value of each visible minority characteristic for the census tracts in the designated neighbourhoods and outputs a table that displays the visible minority in one column and the summed value of the characteristic across all the census tracts in all four neighbourhoods managed by the Southeast Scarborough Planning Table in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?visibleminority (sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137,
        toronto:neighbourhood139 ) )
    ?visibleminority rdfs:subClassOf
    uoft:TotalVisibleMinorityPopulation2016.
    ?visibleminoritygroup a ?visibleminority;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by ?person
    FILTER NOT EXISTS { ?person rdfs:subClassOf uoft:Male.}
    FILTER NOT EXISTS { ?person rdfs:subClassOf uoft:Female.}
}
GROUP BY ?visibleminority

```



Supplementary Figure 6: Results of the visible minority query

	visibleminority	sumvalue
1	uoft:SouthAsianVisibleMinority2016	*37280**xsd:integer
2	uoft:BlackVisibleMinority2016	*18065**xsd:integer
3	uoft:FilipinoVisibleMinority2016	*9235**xsd:integer
4	uoft:ChineseVisibleMinority2016	*5670**xsd:integer
5	uoft:VisibleMinorityNIE2016	*2750**xsd:integer
6	uoft:WestAsianVisibleMinority2016	*2670**xsd:integer
7	uoft:MultipleVisibleMinorities2016	*2475**xsd:integer
8	uoft:LatinAmericanVisibleMinority2016	*1925**xsd:integer
9	uoft:ArabVisibleMinority2016	*1000**xsd:integer
10	uoft:SoutheastAsianVisibleMinority2016	*810**xsd:integer
11	uoft:KoreanVisibleMinority2016	*375**xsd:integer
12	uoft:JapaneseVisibleMinority2016	*240**xsd:integer

Supplementary Figure 7: Results of the senior population query

	neighborhood	sumvalue
1	:#neighborhood137	*8010**xsd:integer
2	:#neighborhood136	*4240**xsd:integer
3	:#neighborhood135	*2695**xsd:integer
4	:#neighborhood139	*2225**xsd:integer

ORDER BY DESC(?sumvalue)

By running the SPARQL query, we can see (Supplementary Figure 6) that the top three largest visible minority groups for the NIA's managed by the Southeast Scarborough Planning Table are South Asian with 37280 residents, Black with 18065 residents, and Filipino with 9235 residents.

CQ8: Which NIAs have the largest senior populations?

The following query finds the value of the characteristic “65+ Years” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood
(sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood oronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?senior a uoft:65+Years2016;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
  ?population a ?populationclass.
  ?populationclass iso21972:defined_by
    uoft:Person65+Years2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 7) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of senior (age 65+) residents (8010) while Neighbourhood 139 (Scarborough Village) has the lowest number of senior residents (2225).

Supplementary Figure 8: Results of the child population query

	neighborhood	sumvalue
1	.#neighborhood137	9625^^xsd:integer
2	.#neighborhood136	4635^^xsd:integer
3	.#neighborhood139	3365^^xsd:integer
4	.#neighborhood135	2880^^xsd:integer

CQ9: Which NIAs have the largest child populations?

The following query finds the value of the characteristic “0-14 Years” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood (sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137, toronto:neighbourhood139 ) )
    ?child a uoft:0-14Years2016;
        uoft:hasLocation ?censustract;
        iso21972:cardinality_of ?population;
        iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by uoft:Person0-14Years2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 8) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of child (age 0-14) residents (9625) while Neighbourhood 135 (Morningside) has the lowest number of child residents (2880).

CQ10: Which NIAs have the longest commute times?

The following query finds the value of the characteristic “commuting duration 60 minutes and over” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137,
        toronto:neighbourhood139 ) )
    ?commute a uoft:60MinutesAndOver2016;
        uoft:hasLocation ?censustract;
        iso21972:cardinality_of ?population;
        iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by
        uoft:Person60MinutesAndOver2016
}
GROUP BY ?neighbourhood

```



Supplementary Figure 9: Results of the commute time query

	neighborhood	sumvalue
1	:#neighborhood137	"5320""xsd:integer
2	:#neighborhood136	"3275""xsd:integer
3	:#neighborhood135	"2010""xsd:integer
4	:#neighborhood139	"1775""xsd:integer

Supplementary Figure 10: Results of the public transit query

	neighborhood	sumvalue
1	:#neighborhood137	"7635""xsd:integer
2	:#neighborhood136	"3990""xsd:integer
3	:#neighborhood139	"2570""xsd:integer
4	:#neighborhood135	"2485""xsd:integer

ORDER BY DESC(?sumvalue)

By running the SPARQL query, we can see (Supplementary Figure 9) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents (5320) with a commute time of 60+ minutes while Neighbourhood 139 (Scarborough Village) has the lowest number of residents (1775) with a commute time of 60+ minutes.

CQ11: Which NIAs use public transit as the main mode of commuting?

The following query finds the value of the characteristic “public transit” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood (sum(?value) as ?sumvalue)
WHERE{
    ?neighbourhood toronto:hasCensusTract ?censustract.
    FILTER (?neighbourhood IN (toronto:neighbourhood135,
        toronto:neighbourhood136, toronto:neighbourhood137,
        toronto:neighbourhood139 ) )
    ?transit a uoft:PublicTransit2016;
    uoft:hasLocation ?censustract;
    iso21972:cardinality_of ?population;
    iso21972:value ?measure.
    ?measure iso21972:numerical_value ?value.
    ?population a ?populationclass.
    ?populationclass iso21972:defined_by
    uoft:PersonPublicTransit2016
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 10) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of residents (7635) who use public transit as their main mode of commuting while Neighbourhood 135 (Morningside) has the lowest number of residents (2485) who use public transit as their main mode of commuting.

CQ12: Which NIAs have the most households living in unsuitable dwellings (according to NOS)?

The following query finds the value of the characteristic “not suitable housing” for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>

```

Supplementary Figure 11: Results of the unsuitable dwelling query

	neighborhood	sumvalue
1	:#neighborhood137	"3535"^^xsd:integer
2	:#neighborhood139	"1320"^^xsd:integer
3	:#neighborhood136	"1220"^^xsd:integer
4	:#neighborhood135	"990"^^xsd:integer

Supplementary Figure 12: Results of the old dwelling query

	neighborhood	sumvalue
1	:#neighborhood137	"6010"^^xsd:integer
2	:#neighborhood136	"2255"^^xsd:integer
3	:#neighborhood139	"1430"^^xsd:integer
4	:#neighborhood135	"1180"^^xsd:integer

```

SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood toronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?notsuitable a uoft:NotSuitable2016;
  uoft:hasLocation ?censustract;
  iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 11) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of households (3535) who live in unsuitable dwellings while Neighbourhood 135 (Morningside) has the lowest number of households (990) who live in unsuitable dwellings.

CQ13: Which NIAs have the highest number old dwellings?

The following query finds the value of the characteristic "1960 or before construction" for the census tracts in the designated neighbourhoods and outputs a table that displays the neighbourhood in one column and the summed value of the characteristic across all the census tracts in that neighbourhood in another column.

```

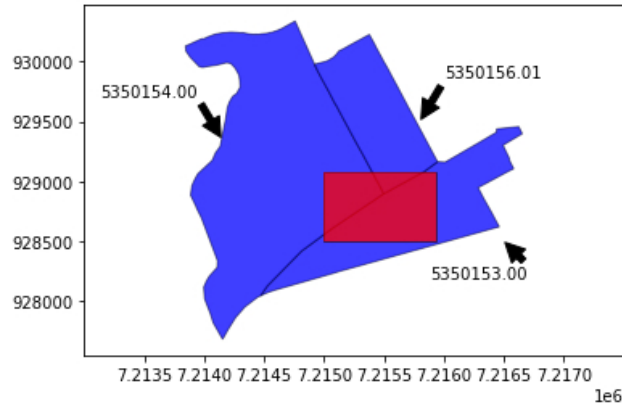
PREFIX uoft: <http://ontology.eil.utoronto.ca/tove/cacensus#>
PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>
PREFIX iso21972: <http://ontology.eil.utoronto.ca/ISO21972/iso21972#>
SELECT ?neighbourhood(sum(?value) as ?sumvalue)
WHERE{
  ?neighbourhood toronto:hasCensusTract ?censustract.
  FILTER (?neighbourhood IN (toronto:neighbourhood135,
    toronto:neighbourhood136, toronto:neighbourhood137,
    toronto:neighbourhood139 ) )
  ?olddwelling a uoft:1960OrBeforeConstruction2016;
  uoft:hasLocation ?censustract;
  iso21972:value ?measure.
  ?measure iso21972:numerical_value ?value.
}
GROUP BY ?neighbourhood
ORDER BY DESC(?sumvalue)

```

By running the SPARQL query, we can see (Supplementary Figure 12) that out of all the NIAs managed by the Southeast Scarborough Planning Table, Neighbourhood 137 (Woburn) has the highest number of dwellings (6010) that were constructed before 1960 while Neighbourhood 135 (Morningside) has the lowest number of dwellings (1180) that were constructed before 1960.



Supplementary Figure 13: Matplotlib visualisation of three census tract geometries (blue) and one non-census geometry (red)



As the Canadian Census Ontology also includes geospatial information, we can use this information to approximate census characteristic values for non-census geometries. Statistics Canada publishes²⁵ cartographic boundary files for census areas in various file formats (e.g., SHP, GML) and these files contain the polygon shapes for representing the geometry of census areas. We can then use functions from Python packages, such as Shapely, to calculate how much an administrative area (e.g., Ward, Neighbourhood) overlaps with the census tract areas.

CQ14: How much of a given geometry overlaps with the census tract boundaries?

Although the SPARQL query language does not have any built-in support for calculating overlapping geometries, programming languages, such as Python, can be used for handling geospatial work. Although there has been increased interest and use of grids and gridded data for aggregating new/future census data into small and uniform land areas, much of the historical census data is published for administrative areas so there may be instances where there is a need for calculating intersecting areas. In Supplementary Figure 13 shown above, a red, non-census geometry polygon overlaps with three different blue census tract polygons. Using the intersection function from the Shapely Python package, we can calculate the area of each overlapping section and divide them by the area of the corresponding census tract polygons in order to obtain the percentages. In this example, the red polygon occupies 34% of the total area of the 5350153.00 census tract, 7% of the total area of the 5350154.00 census tract, and 6% of the total area of the 5350156.01 census tract. With this, we can then multiply the numerical value of the census characteristics for these three census tracts by the percentages in order to obtain an approximate value of the characteristic for the red polygon. For instance, the 5350153.00 census tract has a population of 4462 people, the 5350154.00 census tract has a population of 6157 people, and the 5350156.01 census tract has a population of 5688 people. Multiplying these numbers by their corresponding percentages would give us an approximate population of 2289 for the area represented by the red polygon. Thus, this approach would allow us to approximate values of census characteristics for non-census areas.

²⁵<https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm>.