Navigating the Health PCCF+ Directory

KELLY LEBLANC

UNIVERSITY OF ALBERTA

KLL@UALBERTA.CA

An MLIS Perspective

At Present:

- 2nd year MLIS student at the U of A
- SLIS Cooperative Student working in the Digital Initiatives Unit of UAL
    - Main role: Data Services
- No previous work with PCCF/PCCF+

Going Forward:

- Would like to implement PCCF+ in future research due to its wide-scoping potential

Presentation Overview

- **What is the PCCF+?**

- **Project Development - Navigating the Health PCCF+ Directory**

- **Tools Created:**

  - Health PCCF+ Reference Spreadsheet

  - Health PCCF+ Summary Report

- **Examples & Visuals**

- **Concluding Remarks**

- **What's Next?**

- **References**

(commons.wikimedia.org)

What is the PCCF+?

▶ SAS© control program & set of associated datasets

    ▶ Derived from:

        ▶ Postal Code Conversion File (PCCF)

        ▶ Postal Code population weighted files

        ▶ Geographic Attribute Files

        ▶ Health Region boundary files

    ▶ Automatically assigns a range of StatCan's standard geographic areas & other identifiers based on Postal Codes

    ▶ Uses population-weighted random allocation for Postal Codes

    ▶ Weighted population files link to more than one geographic area.

    ▶ Click Here for the most recent *Postal CodeOM Conversion File Plus (PCCF+) version 6C Reference Guide*

How does the PCCF Differ from the PCCF+?

▶ The Postal Code Conversion File (PCCF) links six-character postal codes to standard geographic areas

  ▶ Through linking, the file facilitates the extraction and subsequent aggregation of data for selected geographic areas.

  ▶ Postal codes are associated with longitude & latitude coordinates to support mapping applications.

▶ PCCF+ complements the PCCF

▶ When the association between postal code & census geography is **not** unique, the PCCF+ allows for a proportional allocation based on weighted population.

(Statistics Canada, 2016)

Navigating the Health PCCF+ Directory

▶ **WHY?**

　　▶ The directory can be difficult to efficiently navigate for new & experienced users

　　▶ Decrease stress for the end user through a smoother data collection process

▶ **WHAT?**

　　▶ Multilevel directory structures can cause a delay in data access -- a **bridge** would promote optimal navigation.

　　▶ The **bridge**:

　　　　▶ (1) The *Health PCCF+ Reference Spreadsheet*

　　　　▶ (2) *Health PCCF+ Summary Report*

▶ **THE ULTIMATE GOAL**

　　▶ Develop tools aimed at assisting users to optimally navigate the Health PCCF+ directory

　　▶ Increase time efficiently & ease of use for users

Directory Folder Name	Folder Type	Folder Descriptor	Guide Type	Guide Title	Guide Location	Guide Doc Name	Guide Format	Postal Code Dates	Release Date	Additional Instructional Docs*	Census Data Year	Postal Code Notes		Notes

Tools Developed

(1) The Reference Spreadsheet

▶ Finding-aid for Health PCCF+ files

▶ Parses out file structure & file types

▶ Provides the location, naming, and details relating to the StatCan Guides.

▶ Answers "where and what" questions

(2) The PCCF+ Summary Report

▶ Consultation manual for the Reference Spreadsheet -- includes visuals!

▶ Outlines & defines content of the Reference Spreadsheet

▶ Provides additional directory navigation guidance

Health PCCF+ Reference Spreadsheet

▶ **What are the key elements?**

    ▶ Folder Name

    ▶ Folder Type

    ▶ StatCan Guide Information – User's Guide/Reference Guide

        ▶ Includes formal Guide title, location within directory, file name, format, release date

    ▶ Postal Code Dates (correspond to the Guide version)

    ▶ Additional Instructional / Informational Documents in directory

    ▶ Census Data Year

    ▶ Postal Code Notes/ Notes – *where applicable*

Health PCCF+ Reference Spreadsheet (cont.)

▶ Total: 14 Directory Folders

▶ Spreadsheet is parsed by **Directory Folder**

▶ **Columns** reveal the data contained within each **directory** listing.

  ▶ *Table 1. List of Column Titles & Descriptions*

    ▶ All columns are mandatory except where noted

    ▶ Add clarity for the novice and experienced user

Table of Column Titles and Corresponding Descriptions

Column Title	Description*
Directory Folder Name	Root folder title
Folder Type	File Folder OR 7z Archive (zip)
Folder Descriptor	Description of folders, file formats, and guides found in the folder
Guide Type	User's Guide (UG) or Reference Guide (RG)
Guide Title	Formal title exhibited on the UG or RG
Guide Location	Location of Guide in the directory
Guide Doc Name	Guide document file name
Guide Format	DOC and/or PDF
Postal Code Dates	Postal code date on cover of guide
Release Date	Release date on guide cover (when available)
Additional Instructional Docs	*(optional)* Additional instructional documents to assist in user clarity (e.g. readme.txt)
Census Data Year	Year (YYYY) of census geography used
Postal Code Notes	*(optional)* Notes related to Postal Code/PCCF/PCCF+
Notes	*(optional)* Notes unrelated to Postal Code

*Column Title fields are mandatory except where noted.

Table 1. Column Titles and Corresponding Descriptions

Health PCCF+ Reference Spreadsheet (cont.)

▶ For Authenticity - source format & naming conventions have not been altered in the following columns:

▶ *Directory Folder Name, Guide Title, Guide Doc Name, Additional Instructional Docs.*

Example taken from the Health PCCF+ Reference Spreadsheet

Directory Folder Name	Guide Title	Guide Doc Name	Additional Instructional Docs*
pccf2j-fccp2j	GEOCODES/PCCF VERSION 2 USER'S GUIDE	geofccp2; geopccf2	\pccf2j-fccp2j\pccf2j.zip\ - README.TXT
pccf3j-fccp3j	PCCF+ VERSION 3J USER'S GUIDE (GEOCODES/PCCF)	msword-fccp3j; msword-pccf3j	\pccf3i-fccp3i - Readpccf3j.txt
pccf4j-fccp4j	PCCF+ Version 4J User's Guide	msword.fccp4j; msword.pccf4j	msword.fmt4jgeo.doc, msword.fmt4jprb.doc

pg. 4 Health PCCF+ Reference Spreadsheet Summary Report

• Naming convention variants in *Guide Doc Name*

  • No hyphen (geofccp2)

  • Hyphen and file type noted (msword-fccp3j)

  • Period and file type noted (msword.pccf4j)

Health PCCF+ Reference Spreadsheet (cont.)
Directory Folder: Contents

▶ *Folder Descriptor* column contains a summary of the directory folder contents.

▶ Contains the type of guide utilized with corresponding language(s), generalized sub-folder/file information *(where applicable)*, file format/extensions used in the directory.

Directory Folder Name	Folder Type	Folder Descriptor
pccf3j-fccp3j	File Folder	Contains User's Guide (eng, fra) file formats: CAN, DOC, GEI, GEO, PDF, PRB, PRI, SAS, TXT
pccf5k-fccp5k-canada-ver3j	7z Archive (.zip)	User's Guide (eng, fra), PRB & GEO doc; file formats: ALL, CAN, DOC, GEO, IGE, IGE3, IPR, IPR3, PDF, PRB, PRBX, SAS, TXT

Type of Guide (User's) & available languages: eng, fra

File format/Extensions: ALL, CAN, DOC, GEO, IGE, IGE$_3$, IPR, IPR$_3$, PDF, PRB, PRBX, SAS, TXT

**File formats are in alphabetical/alphanumeric order for ease of reference

Health PCCF+ Summary Report

Table of Contents   **Page**

Table of Contents from the Health PCCF+ Summary Report

▶ Refers **directly** to Reference Spreadsheet content

    ▶ E.g. *Folder Contents* in the table of contents refers directly to the *Folder Descriptor* column in the Spreadsheet -- and expands upon it.

▶ Visual Examples: Provided in the **Report** content & appendices

Health PCCF+ Summary Report (cont.)

▶ Explanation of timeline related data located in the Spreadsheet (where applicable)

▶ Provides user with reference list of documents & items surveyed for this project

▶ Deliberate redundancy of key information (in both the Report & Spreadsheet) for quick reference purposes, e.g.:

   ▶ Abbreviations Key

   ▶ List of column headers

Abbreviations Key (in order of appearance)	
pccf	- English
fccp	- French
UG	- User's Guide
RG	- Reference Guide
eng	- English
fra	- French

E.g. Abbreviations Key used in both the Spreadsheet & Report

Directory Structure Sample

Directory Listing for Root Folder *pccf2j-fccpf2j*

```
/Health-PCCF-plus-Sante-FCCP-plus
    /pccf2j-fccp2j
        /document-guides
        geofccp2.doc
        geofccp2.pdf
        geopccf2.doc
        geopccf2.pdf
        /pccf2j.zip
        SAMPLDAT.PRB
        SAMPLDAT.IPR
        SAMPLDAT.IGE
        SAMPLDAT.GEO
        SAMPLDAT.CAN
        README.TXT
        PCCFJ97.UNIQ.CAN
        PCCFJ97.RPO.CAN
        PCCFJ97.POINTDUP.CAN
        PCCFJ97.FSAGEOG.CAN
        PCCFJ97.FSA12GEO.CAN
        PCCFJ97.DUPS.CAN
        PCCFJ97.CSDNAMES.CAN
        PCCFJ97.CPCOMM.CAN
        PCCFJ97.CMAIND.CAN
        PCCFJ97.CDNAMES.CAN
        PCCFC91.WCFUDUP3.CAN
        PCCFC91.WCFPOIN3.CAN
        PCCFC91.FSAUDUP2.CAN
        PCCFC91.FSAPOIN2.CAN
        GEORES2J.SAS
        GEORES2J.LST
        GEORES2J.LOG
        GEOPCCF2.pdf
        GEOPCCF2.DOC
        GEOINS2J.SAS
        GEOINS2J.LST
        GEOINS2J.LOG
        GEOFCCP2.pdf
        GEOFCCP2.DOC
        BLDG9606.TXTF1EZ.CAN
        /CNTL
            GEORES2J.SAS
            GEOINS2J.SAS
```

Root Folder

Sub-folder for docs

Example of *User's Guide*: .doc/.pdf files in fra/eng

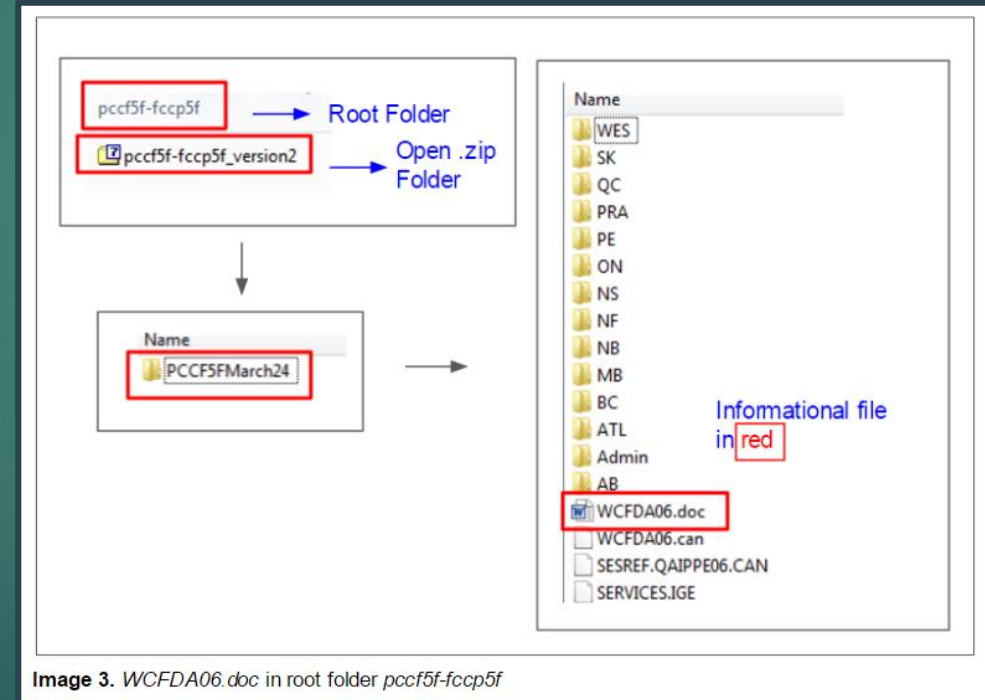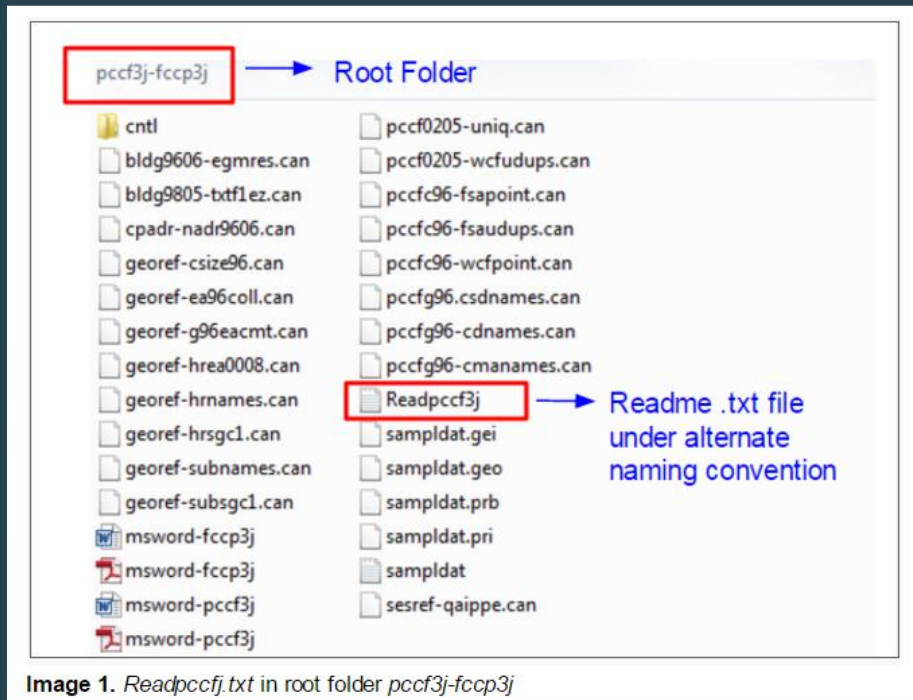Example of *Informational File* located in sub-file /pccf2j.zip

Example of *.SAS* files located in sub-file /CNTL

▶ Directory listing sample (Appendix C) in Summary Report

  ▶ Highlights folder structure

  ▶ Type of file extensions used

  ▶ Naming conventions

  ▶ Variety of informational sources

Informational File Sample

▶ Informational and/or instructional files: not restricted to the naming convention of 'README.txt'

▶ By expanding naming conventions, the Health PCCF+ provides additional informational / instructional files for the benefit of the user.

Image 1. *Readpccfj.txt in root folder pccf3j-fccp3j*

Image 3. *WCFDA06.doc in root folder pccf5f-fccp5f*

Informational File Locations: *Root Folders pccf3j-fccp3j & pccf5f-fccp5f*

Concluding Remarks

- The Health PCCF+ navigation tools are intended to give a generalized overview of the *Health-PCCF-plus* directory.

  - This overview includes (but is not limited to):
    - Directory structure
    - Generalized file folder contents
    - Instructional documents & their corresponding location
  - Both tools aim to assist users with:
    - Troubleshooting the directory structure
    - Answering commonly asked questions

N.B. While the tools created collectively act as comprehensive navigational aids, they are *not* intended to evaluate the contents of the Health PCCF+ Directory.

Potential Next Steps?

▶ Make the tools available to the DLI Community

  ▶ Request feedback

  ▶ Analyze feedback

▶ Determine what the DLI Community wants

▶ Solidify what to provide to the end user

▶ If DLI would like to use the tools:

  ▶ What is the protocol for keeping the tools up to date when new releases come about?

References

Postal CodeOM Conversion File (PCCF), Reference Guide, 2013. Statistics Canada Catalogue no. 92-154-G.

Postal CodeOM Conversion File Plus (PCCF+), Reference Guide, 2014. Statistics Canada Catalogue no. 82-E0086-XDB.

Postal code conversion file (PCCF). (n.d.). Retrieved from
        http://data.library.utoronto.ca/content/postal-code-conversion-file#pccfhlth

Russell Wilkins. *Geocodes/PCCF Version 2 User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion File.* Health Statistics Division, Statistics Canada, Ottawa, July 1997.

Russell Wilkins. *PCCF+ Version 3J User's Guide (Geocodes/PCCF). Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes to May 2002.* Health Analysis and Measurement Group, Statistics Canada, Ottawa, July 2002.

Russell Wilkins. *PCCF+ Version 4J User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through September 2006.* Catalogue 82F0086-XDB. Health Analysis and Measurement Group, Statistics Canada, Ottawa, January 2007.

References (cont.)

Russell Wilkins. *PCCF+ Version 5D User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through September 2008.* Catalogue 82F0086-XDB. Health Information and Research Division, Statistics Canada, Ottawa, March 2009.

Russell Wilkins. *PCCF+ Version 5E User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through March 2009.* Catalogue 82F0086-XDB. Health Analysis Division, Statistics Canada, Ottawa, July 2009.

Russell Wilkins. *PCCF+ Version F User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through July 2009.* Catalogue 82F0086-XDB. Health Analysis Division, Statistics Canada, Ottawa, February 2010.

Russell Wilkins and Saeeda Khan. *PCCF+ Version G User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through December 2009.* Catalogue 82F0086-XDB. Health Analysis Division, Statistics Canada, Ottawa, July 2010.

Russell Wilkins and Saeeda Khan. *PCCF+ Version H User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through October 2010.* Catalogue 82F0086-XDB. Health Analysis Division, Statistics Canada, Ottawa, January 2011.

References (cont.)

Russell Wilkins, Paul A Peters. *PCCF+ Version 5K* User's Guide. Automated Geographic Coding Based on the Statistics Canada Postal Code Conversion Files, Including Postal Codes through May 2011.* Catalogue 82F0086-XDB. Health Analysis Division, Statistics Canada, Ottawa, July 2012.

Statistics Canada. Postal CodeOM Conversion File Plus (PCCF+) Version 6B, Reference Guide. November 2014 Postal Codes. Statistics Canada Catalogue no. 82-E0086-XDB. Ottawa, Minister of Industry, 2015.

Statistics Canada. Postal CodeOM Conversion File Plus (PCCF+) Version 6C, Reference Guide. August 2015 Postal Codes. Statistics Canada Catalogue no. 82-E0086-XDB. Ottawa, Minister of Industry, 2016.

Thank you!

Contact Information:

Kelly LeBlanc

University of Alberta Libraries

Kelly.LeBlanc@ualberta.ca

KLL@ualberta.ca