A scenic view of a river flowing through a forested landscape. The river is dark blue with white rapids and a small waterfall. The surrounding forest is dense with green and yellow trees, suggesting autumn. The sky is blue with some clouds. The text is overlaid on the center of the image.

tidyhydat

Making the case for reproducible workflows in hydrology

Sam Albers

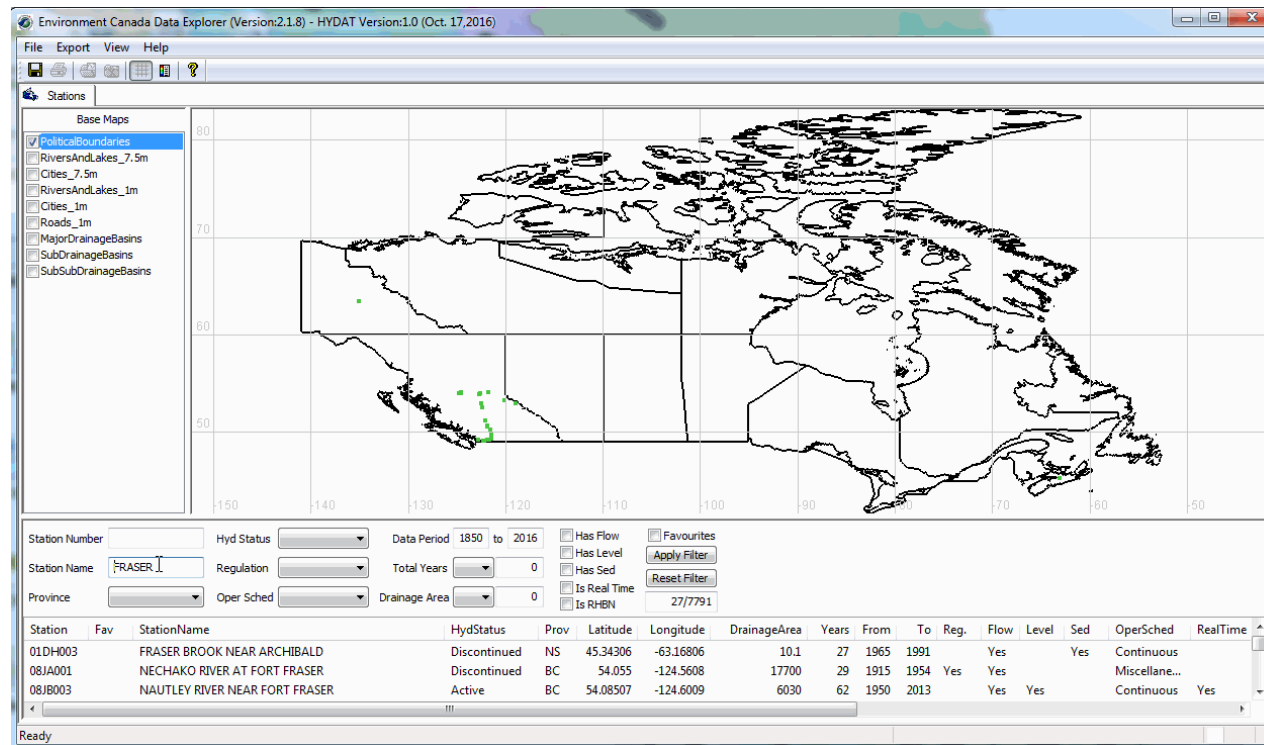
2018-03-01

Outline

- Common Analysis Problems
- What is R and why use it?
- What is tidyhydat?
- Some R basics
- An example of how R can help
- Leveraging R and what I'm not showing you
- Where to get help
- Questions

Common Analysis Problems

Accessing Hydrometric Data



11 clicks!

Stakeholder/Manager: "Hey, this is a really cool analysis but we need to add five stations. Can you run it again?"



Stakeholder/Manager: "Hey, this is a really cool analysis but we need to add five stations. Can you run it again?"



Make it reproducible!

Get off the factory line

How much time do you spend copying and pasting?



Get off the factory line

How much time do you spend copying and pasting?



Automate!

Get off the factory line

How much time do you spend copying and pasting?



Automate!

But how...

...Use R!

(or more generally any programmatic code
based analysis approach...)



What is R?

- Free and open source
- Statistical programming language
- Publication quality graphics
- But definitely not intimidating...

What is R?

- Free and open source
- Statistical programming language
- Publication quality graphics
- But definitely not intimidating...

Why use R?

- Efficient
- Reproducible
- Scalable

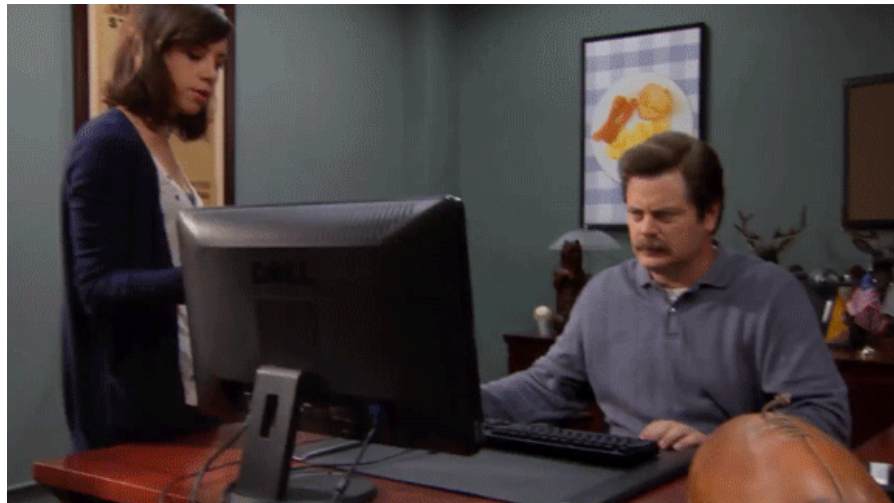
What is R?

- Free and open source
- Statistical programming language
- Publication quality graphics
- But definitely not intimidating...

Why use R?

- Efficient
- Reproducible
- Scalable

Not guaranteed to help with this...



Questions worth asking...

- Are your methods **reproducible**?
- What is your analysis recipe?
- Can you share it?

F

Serves 4

V

Watershed flow correlations

A simple but elegant analysis.

Ingredients

100ml flow data
2 cups tidyhydat
1 cup Butter

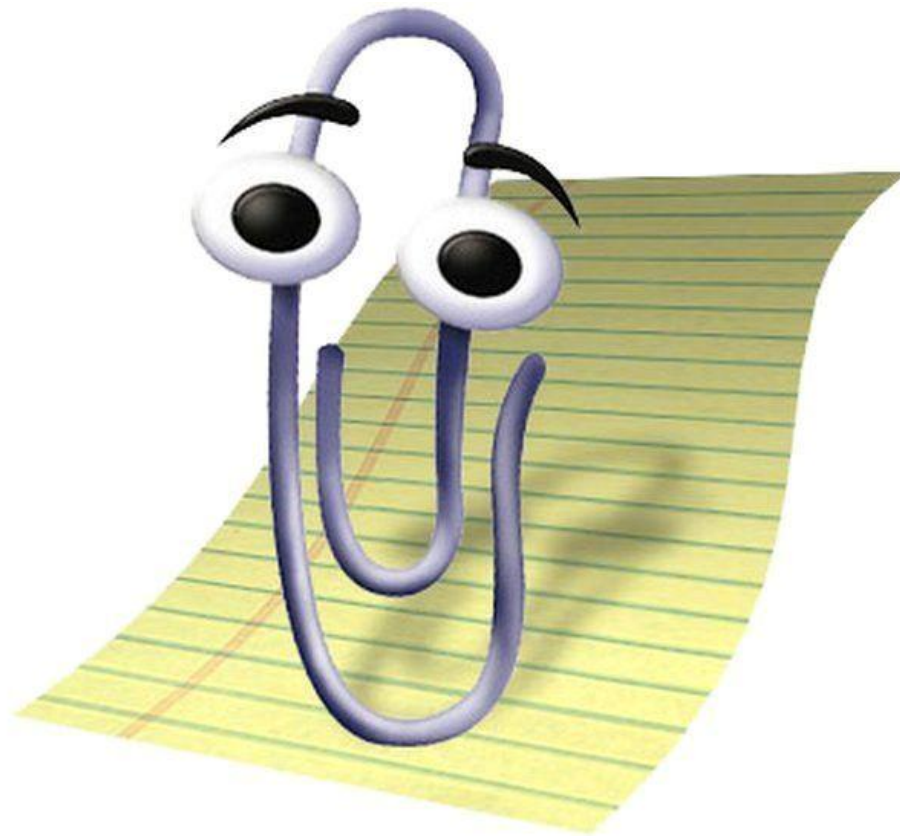
4 tbsp data tidying
1 tsp correlation
3 cups plotted data

Instructions

Preheat the oven to Gas Mark 4, Electric 180°C, Fan 160°C.

1. Stir flow data in a bowl, add tidyhydat and the butter. When the mixture looks like breadcrumbs, mix in the data tidying. Lay the mixture on a shallow baking tray and bake for 25-30 minutes until golden brown. Leave on the side to cool. Mix together the correlation and plotted data and present analysis.

**Excuse me, do you have a moment to
talk about Excel?**

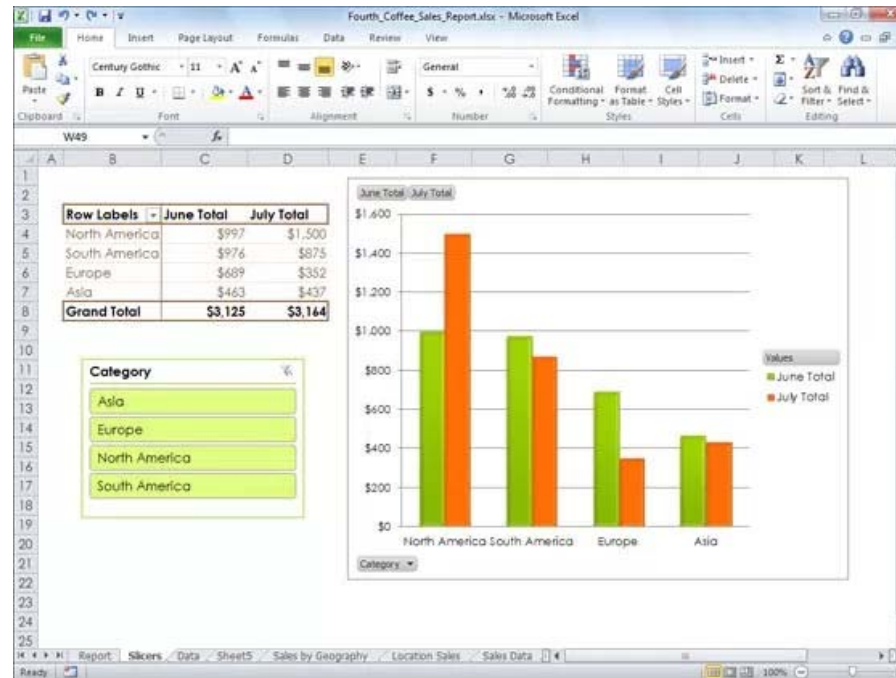


MICROSOFT

R

Excel

Data and analysis are separate Data and analysis are usually stored in the same place



From: <http://blog.yhat.com/posts/R-for-excel-users.html>.

R	Excel
Data and analysis are separate	Data and analysis are usually stored in the same place
Data structure is strict	Data structure is flexible

annex2_data_en [Compatibility Mode] - Microsoft Excel Viewer

Home

Open Quick Print Preview Copy Find Go To Page Setup Print Area Switch Windows Window

Office Document Edit Page Setup Window

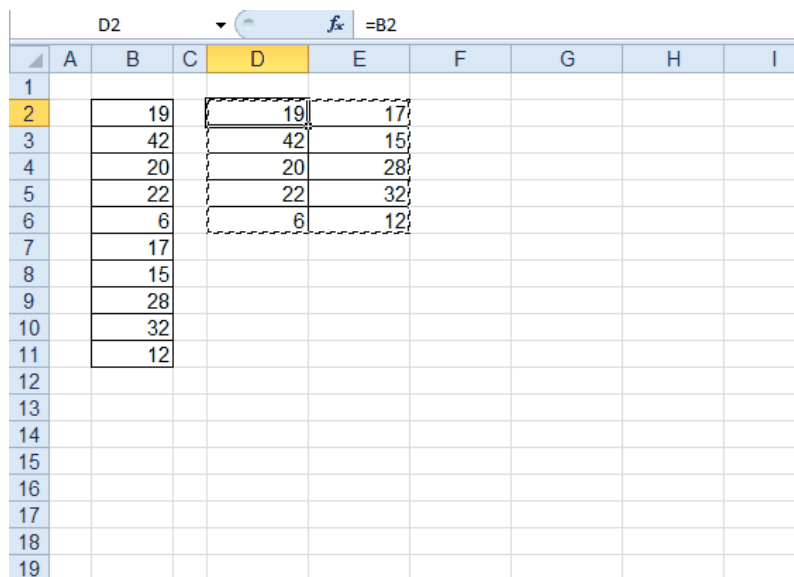
1	Annex 2: HIV and AIDS estimates and data, 2005 and 2003.				
2	Source: 2006 Report on the global AIDS epidemic, UNAIDS/WHO, May 2006.				
3	1. Estimated number of people living with HIV				
4		Adults and children 2005	Adults and children 2003	Adult	
5	Country	Estimate	[low estimate - high estimate]	Estimate	[low estimate - high estimate]
6	Global	38 600 000	[33 400 000 - 46 000 000]	36 200 000	[31 400 000 - 42 900 000]
7	Sub-Saharan Africa	24 500 000	[21 600 000 - 27 400 000]	23 500 000	[20 800 000 - 26 300 000]
8	Angola	320 000	[200 000 - 450 000]	300 000	[190 000 - 430 000]
9	Benin	87 000	[57 000 - 120 000]	90 000	[59 000 - 130 000]
10	Botswana	270 000	[260 000 - 350 000]	260 000	[250 000 - 340 000]
11	Burkina Faso	150 000	[120 000 - 190 000]	150 000	[110 000 - 180 000]
12	Burundi	150 000	[130 000 - 180 000]	140 000	[120 000 - 170 000]
13	Cameroon	510 000	[460 000 - 560 000]	490 000	[450 000 - 540 000]
14	Central African Republic	250 000	[110 000 - 390 000]	240 000	[110 000 - 370 000]
15	Chad	180 000	[88 000 - 300 000]	160 000	[79 000 - 270 000]
16	Comoros	<500	[<1000]	<500	[<1000]
17	Congo	120 000	[75 000 - 160 000]	110 000	[72 000 - 160 000]
18	Côte d'Ivoire	750 000	[470 000 - 1 000 000]	710 000	[440 000 - 950 000]
19	Democratic Republic of Congo	1 000 000	[560 000 - 1 500 000]	840 000	[520 000 - 1 400 000]

2006 Global Report

Ready

From: <http://blog.yhat.com/posts/R-for-excel-users.html>.

R	Excel
Data and analysis are separate	Data and analysis are usually stored in the same place
Data structure is strict	Data structure is flexible
Operations are achieved through scripting	Operations are achieved through pointing and clicking



	A	B	C	D	E	F	G	H	I
1									
2		19		19	17				
3		42		42	15				
4		20		20	28				
5		22		22	32				
6		6		6	12				
7		17							
8		15							
9		28							
10		32							
11		12							
12									
13									
14									
15									
16									
17									
18									
19									

From: <http://blog.yhat.com/posts/R-for-excel-users.html>.

R	Excel
Data and analysis are separate	Data and analysis are usually stored in the same place
Data structure is strict	Data structure is flexible
Operations are achieved through scripting	Operations are achieved through pointing and clicking
Iteration is automated	Iteration is usually done by hand

R provides a clear pathway for efficiency and reproducibility through automation and code

From: <http://blog.yhat.com/posts/R-for-excel-users.html>.

The objective of tidyhydat is to provide a standard method of accessing ECCC hydrometric data sources (historical and real time) using a consistent and easy to use interface that employs tidy data principles within the R project.

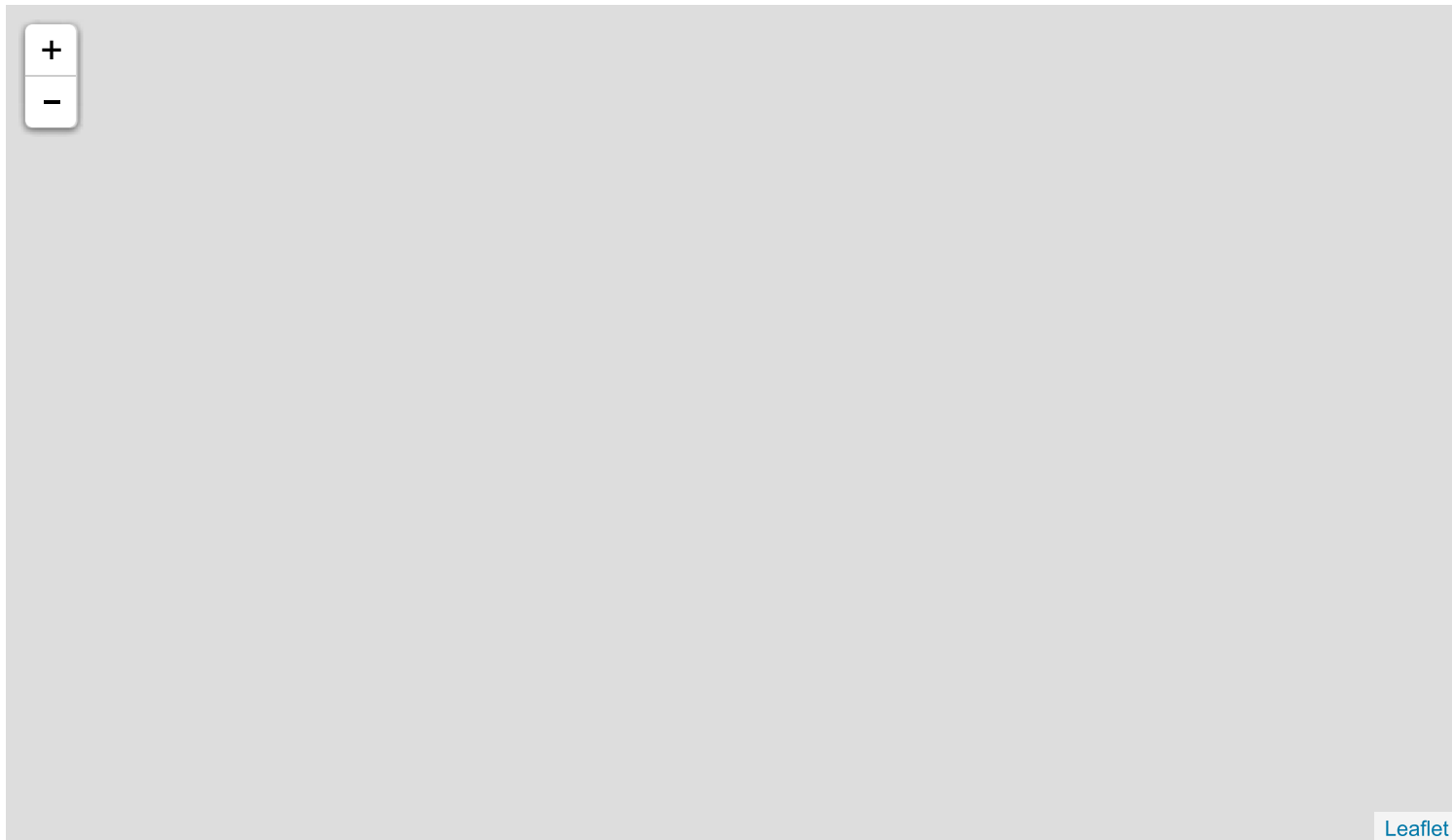


The objective of tidyhydat is to provide a standard method of accessing ECCC hydrometric data sources (historical and real time) using a consistent and easy to use interface that employs tidy data principles within the R project.



tidy|hydat

hydat::Water Survey of Canada Network



tidy::tidy data

■ Tidy datasets are all alike but every messy dataset is messy in its own way¹

tidy::tidy data

Tidy datasets are all alike but every messy dataset is messy in its own way¹

Each variable forms a column

Each observation forms a row



[1] Wickham, Hadley. 2014. Tidy Data. *Journal of Statistical Software* 59 (10). Foundation for Open Access Statistics: 1–23.

tidy::untidy data

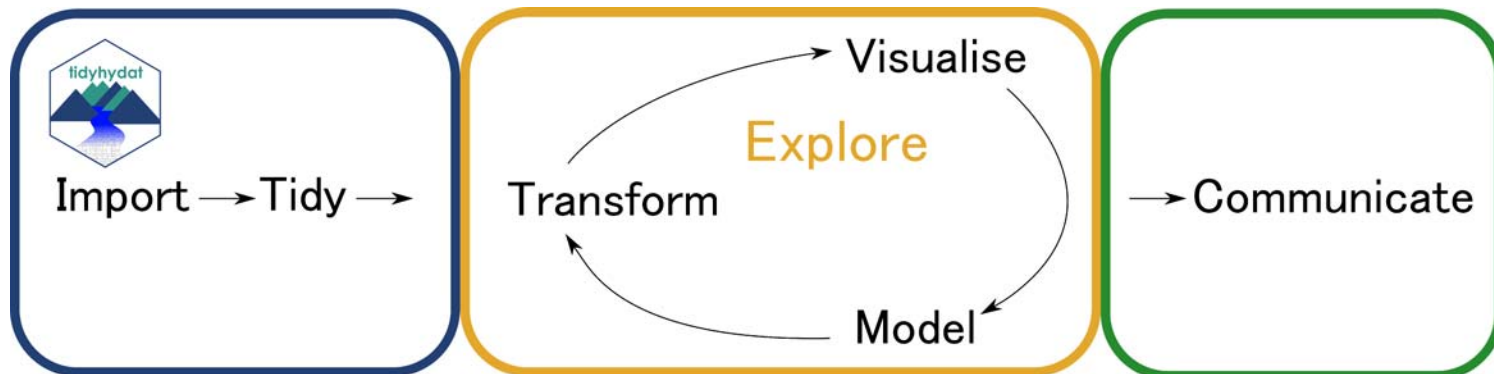
```
#> # Source:   lazy query [?? x 42]
#> # Database: sqlite 3.19.3
#> #   [C:\Users\sallbers\AppData\Local\tidyhydat\tidyhydat\Hydat.sqlite3]
#>   STATION_NUMBER YEAR MONTH FULL_MONTH NO_DAYS MONTHLY_MEAN
#>   <chr>          <int> <int>      <int>    <int>      <dbl>
#> 1 08MF005        1912    3         1       31         485.
#> 2 08MF005        1912    4         1       30        1150.
#> 3 08MF005        1912    5         1       31        4990.
#> 4 08MF005        1912    6         1       30        6130.
#> 5 08MF005        1912    7         1       31        4780.
#> 6 08MF005        1912    8         1       31        3960.
#> 7 08MF005        1912    9         1       30        2160.
#> 8 08MF005        1912   10         1       31        1530.
#> 9 08MF005        1912   11         1       30        1060.
#> 10 08MF005       1912   12         1       31         761.
#> # ... with more rows, and 36 more variables: MONTHLY_TOTAL <dbl>,
#> #   FIRST_DAY_MIN <int>, MIN <dbl>, FIRST_DAY_MAX <int>, MAX <dbl>,
#> #   FLOW1 <dbl>, FLOW2 <dbl>, FLOW3 <dbl>, FLOW4 <dbl>, FLOW5 <dbl>,
#> #   FLOW6 <dbl>, FLOW7 <dbl>, FLOW8 <dbl>, FLOW9 <dbl>, FLOW10 <dbl>,
#> #   FLOW11 <dbl>, FLOW12 <dbl>, FLOW13 <dbl>, FLOW14 <dbl>, FLOW15 <dbl>,
#> #   FLOW16 <dbl>, FLOW17 <dbl>, FLOW18 <dbl>, FLOW19 <dbl>, FLOW20 <dbl>,
#> #   FLOW21 <dbl>, FLOW22 <dbl>, FLOW23 <dbl>, FLOW24 <dbl>, FLOW25 <dbl>,
#> #   FLOW26 <dbl>, FLOW27 <dbl>, FLOW28 <dbl>, FLOW29 <dbl>, FLOW30 <dbl>,
#> #   FLOW31 <dbl>
```

tidy::tidy data

```
#> # A tibble: 37,561 x 5
#>   STATION_NUMBER Date      Parameter Value Symbol
#>   <chr>          <date>    <chr>      <dbl> <chr>
#> 1 08MF005      1912-03-01 Flow        538. <NA>
#> 2 08MF005      1912-03-02 Flow        538. <NA>
#> 3 08MF005      1912-03-03 Flow        538. <NA>
#> 4 08MF005      1912-03-04 Flow        538. <NA>
#> 5 08MF005      1912-03-05 Flow        538. <NA>
#> 6 08MF005      1912-03-06 Flow        538. <NA>
#> 7 08MF005      1912-03-07 Flow        479. <NA>
#> 8 08MF005      1912-03-08 Flow        479. <NA>
#> 9 08MF005      1912-03-09 Flow        459. <NA>
#> 10 08MF005     1912-03-10 Flow        459. <NA>
#> # ... with 37,551 more rows
```

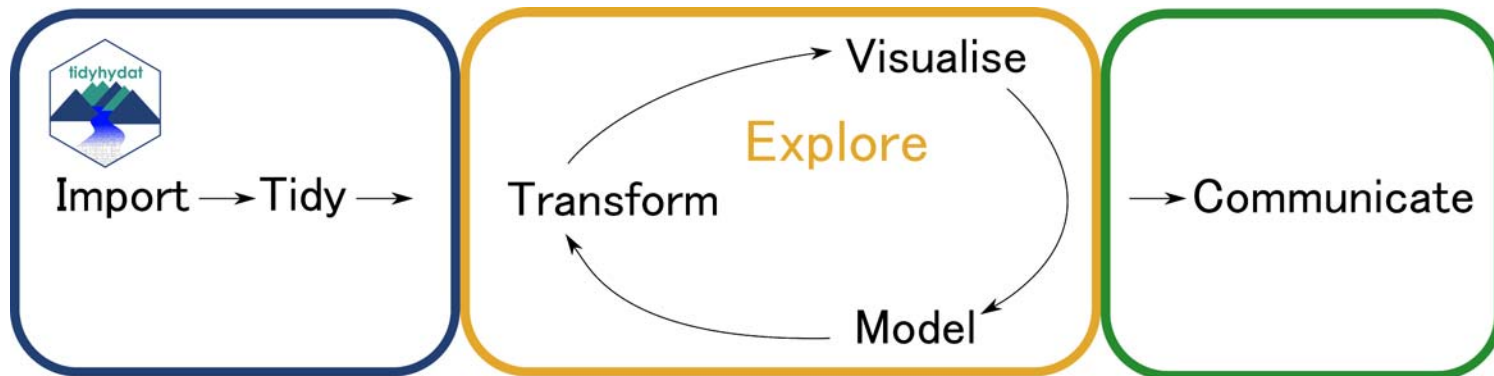


tidy::tidyhydat

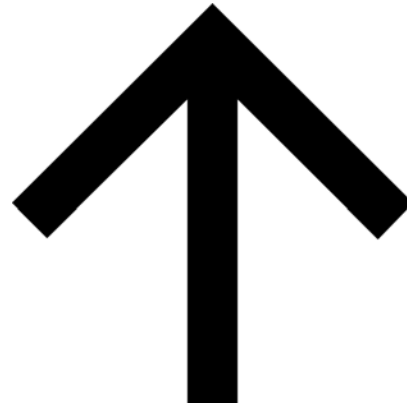


Adapted from Wickham and Grolemund 2017

tidy::tidyhydat



Adapted from Wickham and Grolemund 2017



An Example

tidyhydat & some basic R

```
=SUM(A1:A23)  
=AVERAGE(A1:A23)
```

tidyhydat & some basic R

```
flows_data <- hy_daily_flows(station_number = c("08MF005", "07EF001", "08NE049"))
```

- <=: assignment operator
- flows_data: object
- hy_daily_flows: function
- station_number: argument

tidyhydat & some basic R

```
flows_data <- hy_daily_flows(station_number = c("08MF005", "07EF001", "08NE049"))
```

- <-: assignment operator
- flows_data: object
- hy_daily_flows: function
- station_number: argument

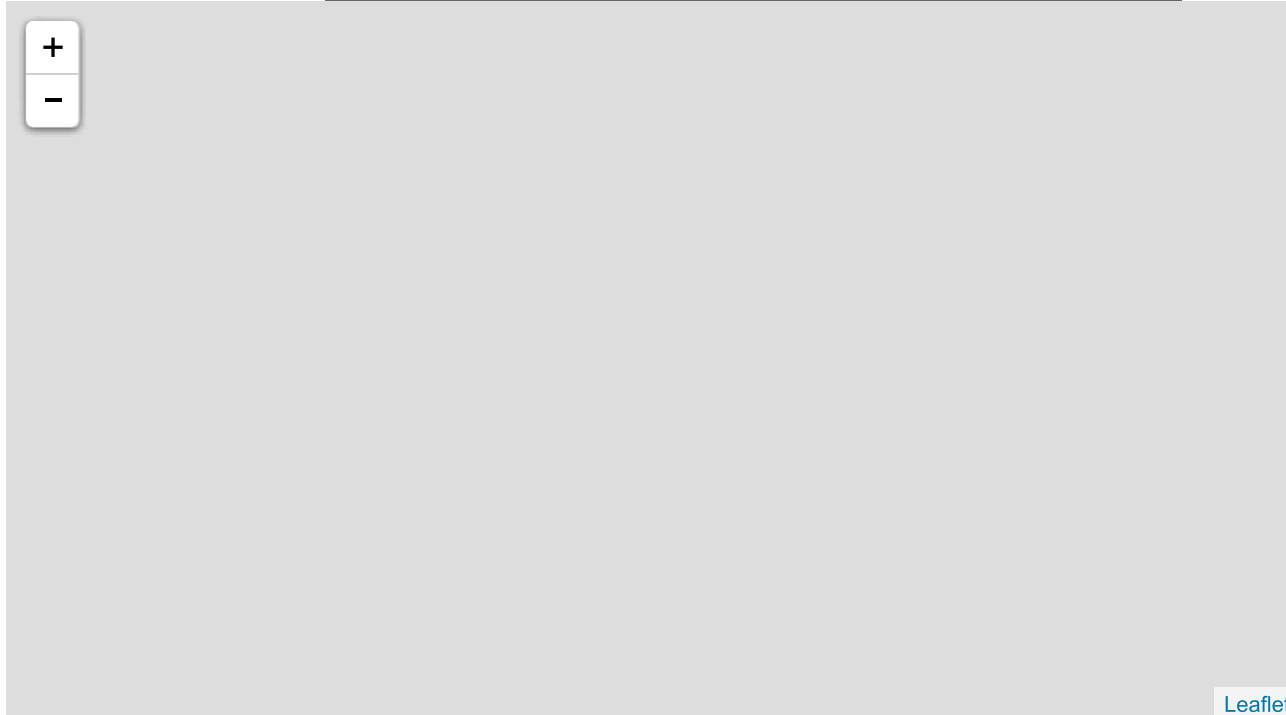
```
flows_data
#> # A tibble: 91,873 x 5
#>   STATION_NUMBER Date      Parameter Value Symbol
#>   <chr>          <date>    <chr>    <dbl> <chr>
#> 1 08MF005      1912-03-01 Flow      538. <NA>
#> 2 08MF005      1912-03-02 Flow      538. <NA>
#> 3 08MF005      1912-03-03 Flow      538. <NA>
#> 4 08MF005      1912-03-04 Flow      538. <NA>
#> 5 08MF005      1912-03-05 Flow      538. <NA>
#> 6 08MF005      1912-03-06 Flow      538. <NA>
#> 7 08MF005      1912-03-07 Flow      479. <NA>
#> 8 08MF005      1912-03-08 Flow      479. <NA>
#> 9 08MF005      1912-03-09 Flow      459. <NA>
#> 10 08MF005     1912-03-10 Flow      459. <NA>
#> # ... with 91,863 more rows
```

Analyze the correlation between:

07EF001 PEACE RIVER AT HUDSON HOPE

08MF005 FRASER RIVER AT HOPE

08NE049 COLUMBIA RIVER AT BIRCHBANK



Build the analysis

```
flows_data
#> # A tibble: 91,873 x 5
#>   STATION_NUMBER Date      Parameter Value Symbol
#>   <chr>          <date>    <chr>    <dbl> <chr>
#> 1 08MF005      1912-03-01 Flow      538. <NA>
#> 2 08MF005      1912-03-02 Flow      538. <NA>
#> 3 08MF005      1912-03-03 Flow      538. <NA>
#> 4 08MF005      1912-03-04 Flow      538. <NA>
#> 5 08MF005      1912-03-05 Flow      538. <NA>
#> 6 08MF005      1912-03-06 Flow      538. <NA>
#> 7 08MF005      1912-03-07 Flow      479. <NA>
#> 8 08MF005      1912-03-08 Flow      479. <NA>
#> 9 08MF005      1912-03-09 Flow      459. <NA>
#> 10 08MF005     1912-03-10 Flow      459. <NA>
#> # ... with 91,863 more rows
```

flows_data: object

Build the analysis

```
flows_data %>%  
  spread(key = STATION_NUMBER, value = Value)  
#> # A tibble: 43,814 x 6  
#>   Date      Parameter Symbol `07EF001` `08MF005` `08NE049`  
#>   <date>    <chr>      <chr>      <dbl>      <dbl>      <dbl>  
#> 1 1912-03-01 Flow      <NA>      NA        538.      NA  
#> 2 1912-03-02 Flow      <NA>      NA        538.      NA  
#> 3 1912-03-03 Flow      <NA>      NA        538.      NA  
#> 4 1912-03-04 Flow      <NA>      NA        538.      NA  
#> 5 1912-03-05 Flow      <NA>      NA        538.      NA  
#> 6 1912-03-06 Flow      <NA>      NA        538.      NA  
#> 7 1912-03-07 Flow      <NA>      NA        479.      NA  
#> 8 1912-03-08 Flow      <NA>      NA        479.      NA  
#> 9 1912-03-09 Flow      <NA>      NA        459.      NA  
#> 10 1912-03-10 Flow     <NA>      NA        459.      NA  
#> # ... with 43,804 more rows
```

%>%: "then"

spread: function

Build the analysis

```
flows_data %>%  
  spread(key = STATION_NUMBER, value = Value) %>%  
  select(-Date, -Symbol, -Parameter)  
#> # A tibble: 43,814 x 3  
#>   `07EF001` `08MF005` `08NE049`  
#>   <dbl>     <dbl>     <dbl>  
#> 1      NA      538.      NA  
#> 2      NA      538.      NA  
#> 3      NA      538.      NA  
#> 4      NA      538.      NA  
#> 5      NA      538.      NA  
#> 6      NA      538.      NA  
#> 7      NA      479.      NA  
#> 8      NA      479.      NA  
#> 9      NA      459.      NA  
#> 10     NA      459.      NA  
#> # ... with 43,804 more rows
```

select: function

Build the analysis

```
flows_data %>%  
  spread(key = STATION_NUMBER, value = Value) %>%  
  select(-Date, -Symbol, -Parameter) %>%  
  correlate()  
#> # A tibble: 3 x 4  
#>   rowname `07EF001` `08MF005` `08NE049`  
#>   <chr>      <dbl>      <dbl>      <dbl>  
#> 1 07EF001    NA        0.297      0.507  
#> 2 08MF005    0.297      NA        0.746  
#> 3 08NE049    0.507      0.746      NA
```

correlation: function

Build the analysis

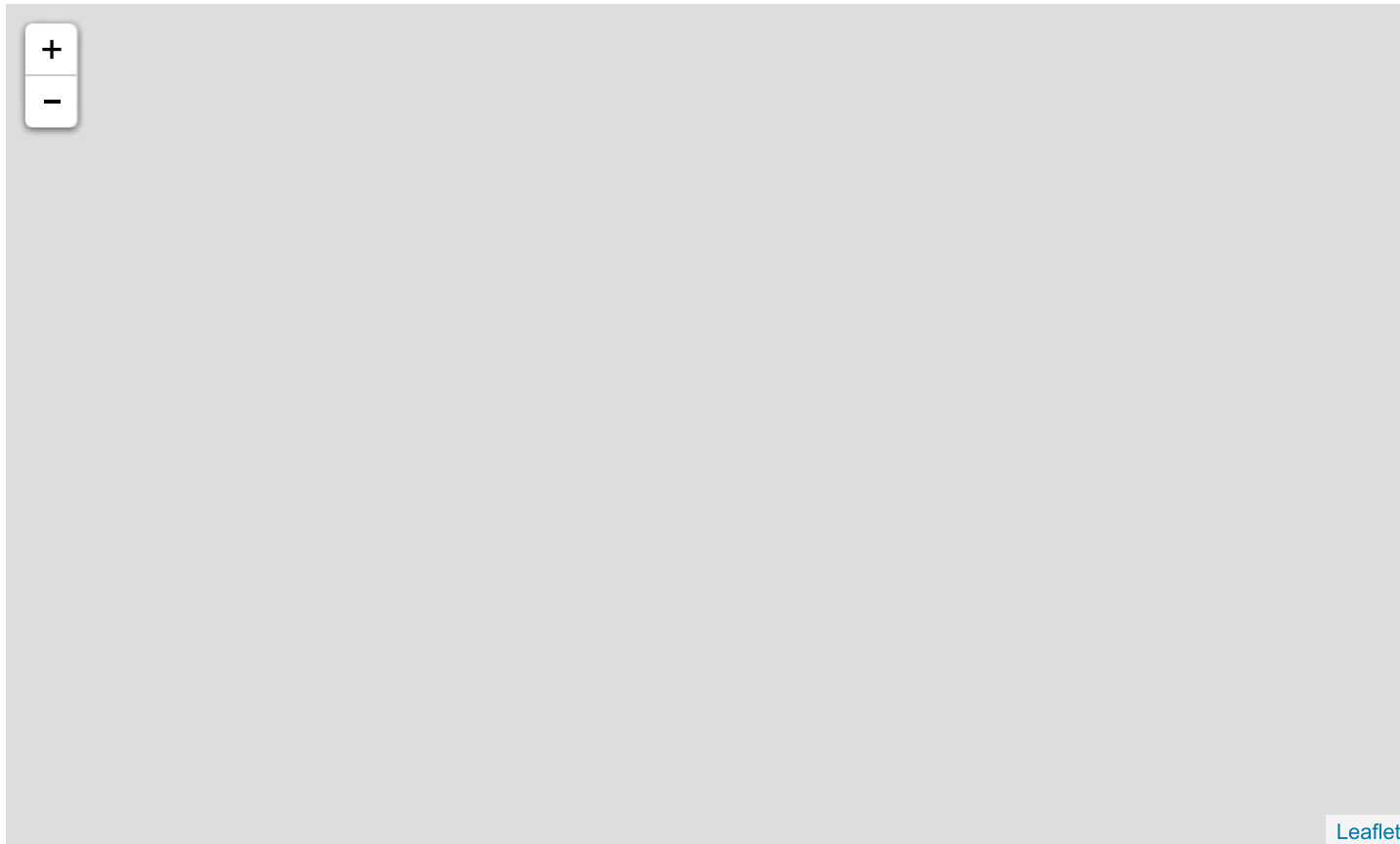
```
flows_data %>%  
  spread(key = STATION_NUMBER, value = Value) %>%  
  select(-Date, -Symbol, -Parameter) %>%  
  correlate() %>%
```

```
stretch()
```

```
#> # A tibble: 9 x 3  
#>   x          y          r  
#>   <chr>    <chr>    <dbl>  
#> 1 07EF001 07EF001  NA  
#> 2 07EF001 08MF005  0.297  
#> 3 07EF001 08NE049  0.507  
#> 4 08MF005 07EF001  0.297  
#> 5 08MF005 08MF005  NA  
#> 6 08MF005 08NE049  0.746  
#> 7 08NE049 07EF001  0.507  
#> 8 08NE049 08MF005  0.746  
#> 9 08NE049 08NE049  NA
```

stretch: **function**

Scalable



Scalable

```
south_vi_stns <- c("08HA001", "08HA002", "08HA003", "08HA009", "08HA010", "08HA011",  
"08HA016", "08HA068", "08HA069", "08HA070", "08HA072")
```

```
south_vi_flows <- hy_daily_flows(station_number = south_vi_stns)
```

```
south_vi_flows
```

```
#> # A tibble: 164,859 x 5
```

```
#>   STATION_NUMBER Date      Parameter Value Symbol
```

```
#>   <chr>          <date>    <chr>    <dbl> <chr>
```

```
#> 1 08HA002      1913-03-01 Flow      49.3 <NA>
```

```
#> 2 08HA002      1913-03-02 Flow      47.3 <NA>
```

```
#> 3 08HA002      1913-03-03 Flow      45.3 <NA>
```

```
#> 4 08HA002      1913-03-04 Flow      43.3 <NA>
```

```
#> 5 08HA002      1913-03-05 Flow      43.3 <NA>
```

```
#> 6 08HA002      1913-03-06 Flow      41.3 <NA>
```

```
#> 7 08HA002      1913-03-07 Flow      41.3 <NA>
```

```
#> 8 08HA002      1913-03-08 Flow      39.6 <NA>
```

```
#> 9 08HA002      1913-03-09 Flow      39.6 <NA>
```

```
#> 10 08HA002     1913-03-10 Flow      39.6 <NA>
```

```
#> # ... with 164,849 more rows
```

Scalable

```
south_vi_flows %>%
  spread(STATION_NUMBER, Value) %>%
  select(-Date, -Symbol, -Parameter) %>%
  correlate() %>%
  stretch()
#> # A tibble: 100 x 3
#>   x           y           r
#>   <chr>    <chr>    <dbl>
#> 1 08HA001 08HA001  NA
#> 2 08HA001 08HA002  0.630
#> 3 08HA001 08HA003  0.860
#> 4 08HA001 08HA010  0.851
#> 5 08HA001 08HA011  0.713
#> 6 08HA001 08HA016  0.768
#> 7 08HA001 08HA068  0.814
#> 8 08HA001 08HA069  0.767
#> 9 08HA001 08HA070  0.882
#> 10 08HA001 08HA072  0.862
#> # ... with 90 more rows
```

What else is available in tidyhydat?

All tables in HYDAT

- Instantaneous peaks
- Daily, monthly and yearly temporal aggregations
- Discharge, level, sediment, particle size
- Data ranges
- Station metadata

What else is available in tidyhydat?

```
search_stn_name("fraser")
#> # A tibble: 32 x 5
#>   STATION_NUMBER STATION_NAME      PROV_TERR_STATE_~ LATITUDE LONGITUDE
#>   <chr>          <chr>          <chr>          <dbl>    <dbl>
#> 1 08JB003      NAUTLEY RIVER NEAR~ BC          54.1     -125.
#> 2 08KA004      FRASER RIVER AT HA~ BC          54.1     -122.
#> 3 08KA005      FRASER RIVER AT MC~ BC          53.3     -120.
#> 4 08KA007      FRASER RIVER AT RE~ BC          53.0     -119.
#> 5 08KB001      FRASER RIVER AT SH~ BC          54.0     -123.
#> 6 08KE018      FRASER RIVER AT SO~ BC          53.9     -123.
#> 7 08MC018      FRASER RIVER NEAR ~ BC          52.5     -122.
#> 8 08MF005      FRASER RIVER AT HO~ BC          49.4     -121.
#> 9 08MF035      FRASER RIVER NEAR ~ BC          49.2     -122.
#> 10 08MF038     FRASER RIVER AT CA~ BC          49.1     -122.
#> # ... with 22 more rows
```

Pointing and clicking


The screenshot shows a web browser window with the URL `https://wateroffice.ec.gc.ca/search/real_time_e.html`. The page features the Government of Canada header with the Canadian flag, the text "Government of Canada" and "Gouvernement du Canada", and a search bar labeled "Search Canada.ca". Below the header is a navigation menu with links for Jobs, Immigration, Travel, Business, Benefits, Health, Taxes, and More services. The main content area has a breadcrumb trail: [Home](#) → [Environment and natural resources](#) → [Water level and flow](#) → [Real-Time Data](#). The title "Real-Time Hydrometric Data Text Search" is followed by instructions: "Search for a station by entering all or part of a station name or station number. Multiple station numbers (up to 20) can also be entered as comma-separated values in the form of '99AA999,99AAA99,etc'." A "Search" button is located below the instructions. The search form has three radio buttons: "Station Name" (selected), "Station Number", and "Province or Territory". The "Station Name" field contains the placeholder text "Enter Full or Partial Station Name". The "Station Number" field contains the placeholder text "Enter Full or Partial Station Number". The "Province or Territory" field is a dropdown menu showing "Alberta".

Real-Time Hydrometric

Secure | https://wateroffice.ec.gc.ca/search/real_time_e.html

GH_R CRAN_R R_examples Sharepoint Docs WSC Aquarius Records_management GOVT_sites Water_Info_Tools Training Delineation

Français

 Government of Canada / Gouvernement du Canada

Search Canada.ca

Jobs Immigration Travel Business Benefits Health Taxes More services

[Home](#) → [Environment and natural resources](#) → [Water level and flow](#) → [Real-Time Data](#)

Real-Time Hydrometric Data Text Search

Search for a station by entering all or part of a station name or station number. Multiple station numbers (up to 20) can also be entered as comma-separated values in the form of "99AA999,99AAA99,etc".

Search

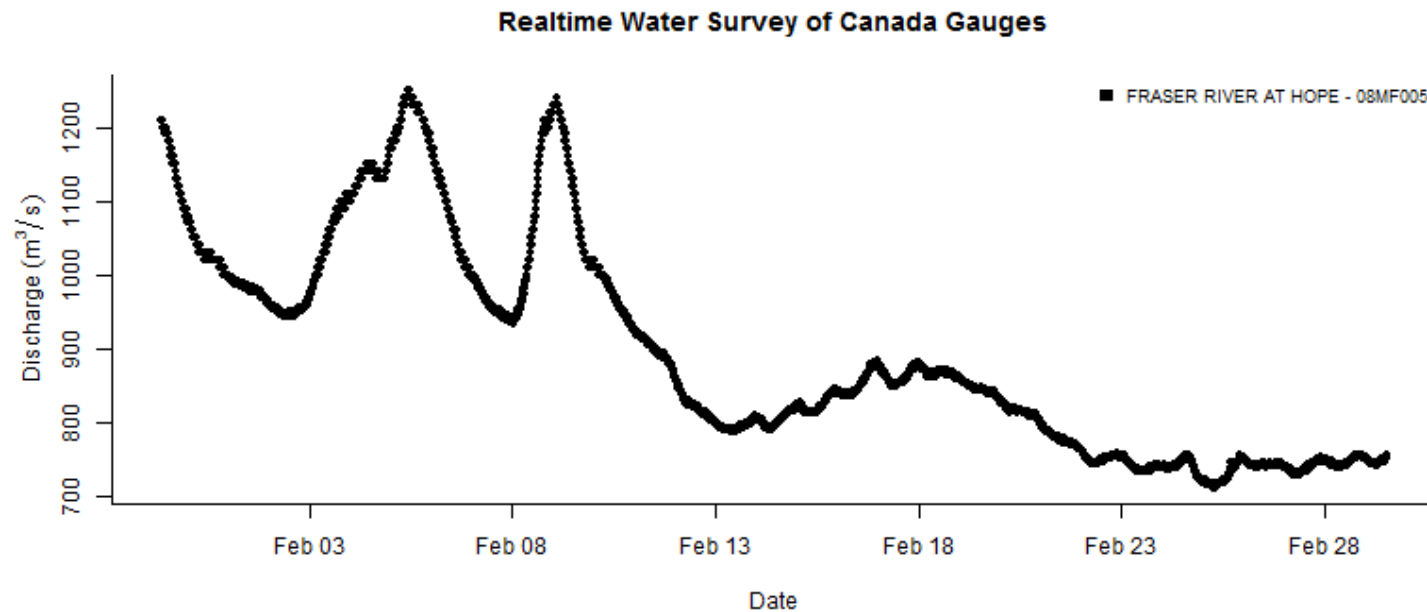
☒ **Station Name** Enter Full or Partial Station Name

☐ **Station Number** Enter Full or Partial Station Number

☐ **Province or Territory** Alberta

What else is available in tidyhydat?

```
realtime_plot("08MF005", Parameter = "Flow")
```



What else is available in R?

```
raw_stns <- hy_stations() %>%
  select(STATION_NUMBER:PROV_TERR_STATE_LOC, DRAINAGE_AREA_GROSS)

mad_long_avg <- hy_annual_stats(raw_stns$STATION_NUMBER) %>%
  filter(Sum_stat == "MEAN", Parameter == "Flow") %>%
  group_by(STATION_NUMBER) %>%
  summarise(Value = mean(Value, na.rm = TRUE)) %>%
  right_join(raw_stns)
```

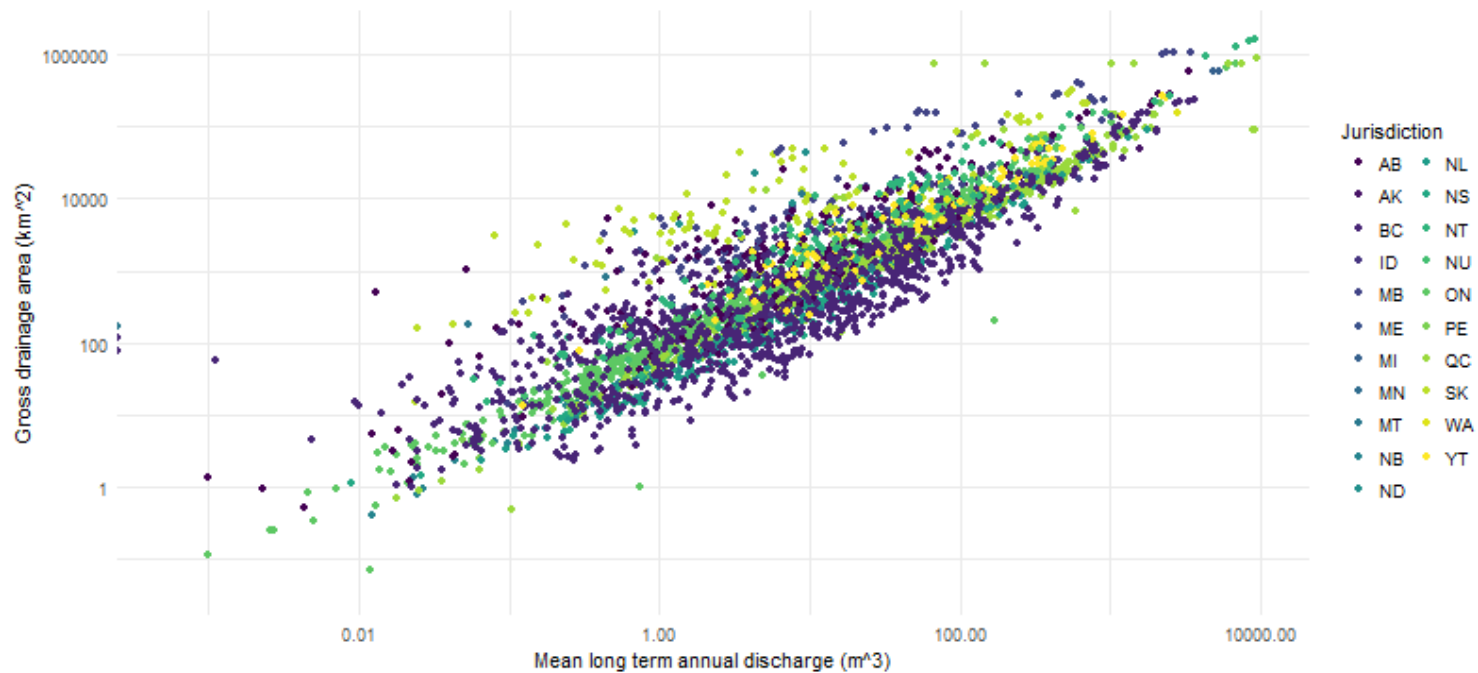
mad_long_avg

```
#> # A tibble: 7,794 x 5
```

```
#>   STATION_NUMBER Value STATION_NAME PROV_TERR_STATE~ DRAINAGE_AREA_G~
#>   <chr>          <dbl> <chr>          <chr>          <dbl>
#> 1 01AA002        17.8 DAAQUAM (RIVIE~ QC          598.
#> 2 01AD001        48.9 MADAWASKA (RIV~ QC          2690.
#> 3 01AD002       279. SAINT JOHN RIV~ ME          14700.
#> 4 01AD003       25.4 ST. FRANCIS RI~ NB          1350.
#> 5 01AD004       328. SAINT JOHN RIV~ NB          15500.
#> 6 01AD005        NA MADAWASKA (RIV~ QC          2590.
#> 7 01AD008        NA LONG (LAC) PRE~ QC           108.
#> 8 01AD009        NA CABANO (RIVIER~ QC           172.
#> 9 01AD012        NA SAINT-FRANCOIS~ QC           58.6
#> 10 01AD013       1.31 SAINT-FRANCOIS~ QC           59.8
#> # ... with 7,784 more rows
```

What else is available in R?

```
library(ggplot2)
ggplot(mad_long_avg, aes(x = Value, y = DRAINAGE_AREA_GROSS, colour = PROV_TERR_STATE_LOC)) +
  geom_point() +
  scale_y_continuous(trans = "log10") +
  scale_x_continuous(trans = "log10") +
  scale_colour_viridis_d(name = "Jurisdiction") +
  labs(x = "Mean long term annual discharge (m^3)", y = "Gross drainage area (km^2)") +
  theme_minimal()
```



It can be daunting!



Resources for R



Contribute to tidyhydat

Openly developed on GitHub 

<https://github.com/ropensci/tidyhydat>

Any contribution helps. You don't have to be an R programmer!

- Questions
- Ideas / Feature-requests
- Bugs
- Bug-fixes
- Development



Contribute to tidyhydat

Openly developed on GitHub

<https://github.com/ropensci/tidyhydat>

Any contribution helps. You don't have to be an R programmer!

- Questions
- Ideas / Feature-requests
- Bugs
- Bug-fixes
- Development



For example...

```
Authors@R: c(person("Sam", "Albers", email = "sam.albers@gov.bc.ca", role = c("aut", "cre")),  
  person("David", "Hutchinson", email = "david.hutchinson@canada.ca", role = "ctb"),  
  person("Dewey", "Dunnington", email = "dewey@fishandwhistle.net", role = "ctb"),  
  person("Province of British Columbia", role = "cph"))
```

Some Helpful Links

Installing R & RStudio on a BC Gov Computer

-<https://github.com/bcgov/bcgov-data-science-resources/wiki/Installing-R-&-RStudio>

Installing tidyhydat

-<https://cran.rstudio.com/web/packages/tidyhydat/README.html>

Getting started with tidyhydat

-https://cran.rstudio.com/web/packages/tidyhydat/vignettes/tidyhydat_an_introduction.html
-https://cran.rstudio.com/web/packages/tidyhydat/vignettes/tidyhydat_example_analysis.html

BC Gov data science resource wiki

-<https://github.com/bcgov/bcgov-data-science-resources/wiki>

Questions?

Slides available from

-https://github.com/ropensci/tidyhydat/blob/master/presentations/tidyhydat_intro.Rmd

Contact sam.albers@gov.bc.ca