Access WAI database using R (Voss PI)

Nicholas Horton (nhorton@amherst.edu)

May 29, 2018

This file provides an example of the use of SQL within R to access data from the Wideband Acoustic Immittance database. See http://www.science.smith.edu/wai-database/ for more information.

```
library(mosaic)
library(RMySQL)
```

```
## Loading required package: DBI
```

```
db <- src_mysql(dbname = "wai", host = "scidb.smith.edu", user = "waiuser",
          password="smith_waiDB")
Measurements <- tbl(db, "Measurements")
PI_Info <- tbl(db, "PI_Info")
Subject <- tbl(db, "Subject")
```

Let's explore the `PI_Info` table.

```
PI_Info %>% collect() %>% data.frame()
```

```
##      Identifier PI_Year                                          PI
## 1 Rosowski_2012    2012                          John J. Rosowski
## 2     Abur_2014    2014 Defne Abur; Nicholas J. Horton; Susan E. Voss
## 3  Shahnaz_2006    2006                Navid Shahnaz; Karin Bork
## 4     Voss_1994    1994                           Susan E. Voss
## 5   Werner_2010    2010                           Douglas Keefe
## 6      Sun_2016    2016                           Xiao-Ming Sun
##
## 1 Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston; Department of Otology and La
## 2
## 3
## 4
## 5
## 6
##                            Email
## 1 John_Rosowski@meei.harvard.edu
## 2                svoss@smith.edu
## 3   nshahnaz@audiospeech.ubc.ca
## 4                svoss@smith.edu
## 5     Douglas.Keefe@boystown.org
## 6      xiao-ming.sun@wichita.edu
##
## 1                                  Ear-Canal Reflectance, Umbo Velocity, and Tympanometry in Normal-
## 2                                                        Intrasubject Variability in Pou
## 3                                          Wideband Reflectance Norms for Caucasian and Chines
## 4                                    Measurement of acoustic impedance and reflectance in the I
## 5                      Ear-Canal Wideband Acoustic Transfer Functions of Adults and Two- to Nine-Mon
## 6 Wideband acoustic immittance: Normative study and test-retest reliability of tympanometric measurem
##                             Pub       Date
## 1                 Ear & Hearing 11/06/2015
```

```
## 2              J. Am Acad Audiol 08/24/2016
## 3                 Ear & Hearing  08/24/2016
## 4 Journal of the Acoustical Soci 02/16/2017
## 5                Ear and Hearing   9/1/2017
## 6 Journal of Speech, Language, a 10/31/2017
##
## 1                                                                                 http://www.ncbi.nlm.nih.gov/pubr
## 2                                                                                                              Nc
## 3  http://journals.lww.com/ear-hearing/Abstract/2006/12000/Wideband_Reflectance_Norms_for_Caucasian_a
## 4                                                                                 http://www.ncbi.nlm.nih.gov/pubr
## 5                                                                                https://www.ncbi.nlm.nih.gov/pubr
## 6                                                                                https://www.ncbi.nlm.nih.gov/pubr
##
## 1 HearID (Mimosa Acoustics); \nNormal Criteria as follows: \n(1) There was no history of significant
## 2
## 3
## 4
## 5
## 6
```

Let's explore the `Subjects` table.

```
Subject %>% collect() %>% data.frame() %>% head()
```

```
##      Identifier Sub_Number Session_Total Age Female Race Ethnicity
## 1 Rosowski_2012          3             1  30      1    5         2
## 2 Rosowski_2012          6             1  29      0    5         2
## 3 Rosowski_2012         11             1  64      1    5         2
## 4 Rosowski_2012         12             1  42      1    5         2
## 5 Rosowski_2012         14             1  24      0    5         2
## 6 Rosowski_2012         15             1  32      1    5         2
##   Left_Ear_Status Right_Ear_Status Left_Ear_Area Right_Ear_Area Sub_Notes
## 1               0                0           NA             NA       NaN
## 2               0                0           NA             NA       NaN
## 3               0                0           NA             NA       NaN
## 4               0                0           NA             NA       NaN
## 5               0                0           NA             NA       NaN
## 6               0                0           NA             NA       NaN
```

Let's explore the `Measurements` table.

```
Measurements %>% summarise(total = n())
```

```
## # Source:   lazy query [?? x 1]
## # Database: mysql 5.5.58-0ubuntu0.14.04.1-log
## #   [waiuser@scidb.smith.edu:/wai]
##    total
##    <dbl>
## 1 131602
```

Let's download the data from a given subject

```
Rosowski <-
  Measurements %>%
```

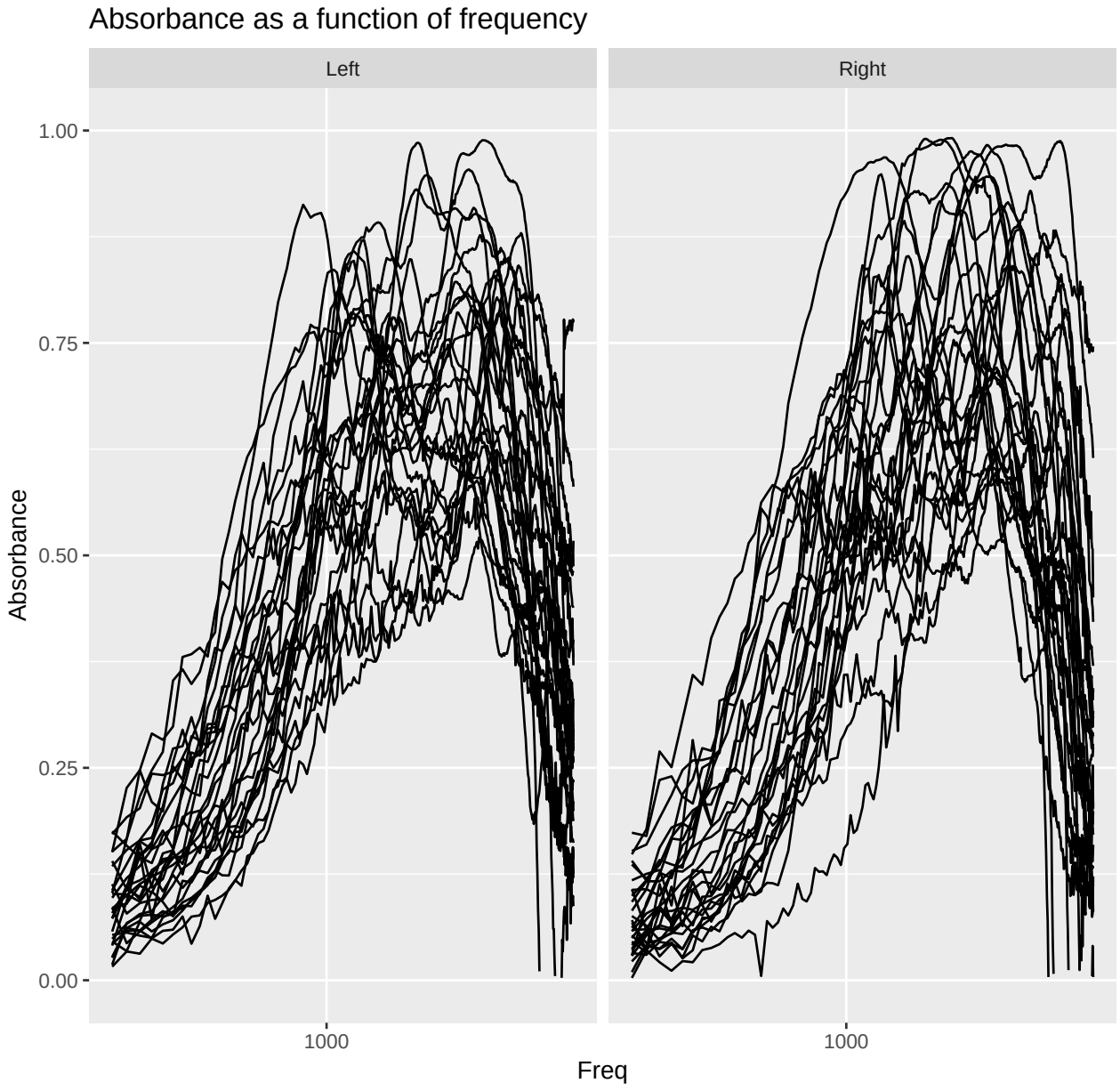
```
  filter(Identifier=="Rosowski_2012") %>%
  collect %>%
  mutate(SessionNum = as.factor(Session),
         EarStatus = ifelse(Left_Ear==1, "Left", "Right")) %>%
  arrange(Sub_Number, Freq, EarStatus)
head(Rosowski)
```

```
## # A tibble: 6 x 12
##   Identifier Sub_Number Session Left_Ear   MEP Instrument  Freq Absorbance
##   <chr>           <int>   <int>    <int> <dbl>      <int> <dbl>      <dbl>
## 1 Rosowski_~          3       1        1    NA          1  211.     0.0852
## 2 Rosowski_~          3       1        0    NA          1  211.     0.0528
## 3 Rosowski_~          3       1        1    NA          1  234.     0.0903
## 4 Rosowski_~          3       1        0    NA          1  234.     0.0365
## 5 Rosowski_~          3       1        1    NA          1  258.     0.112
## 6 Rosowski_~          3       1        0    NA          1  258.     0.0494
## # ... with 4 more variables: Zmag <dbl>, Zang <dbl>, SessionNum <fct>,
## #   EarStatus <chr>
```

and plot the results

```
ggplot(Rosowski, aes(x=Freq, y=Absorbance, group=Sub_Number)) +
  geom_line() +
  facet_grid(. ~ EarStatus) +
  scale_x_log10() +
  ylim(0, 1) +
  labs(title="Absorbance as a function of frequency", caption = "Data from Rosowski et al, 2012")
```

```
## Warning: Removed 66 rows containing missing values (geom_path).
```

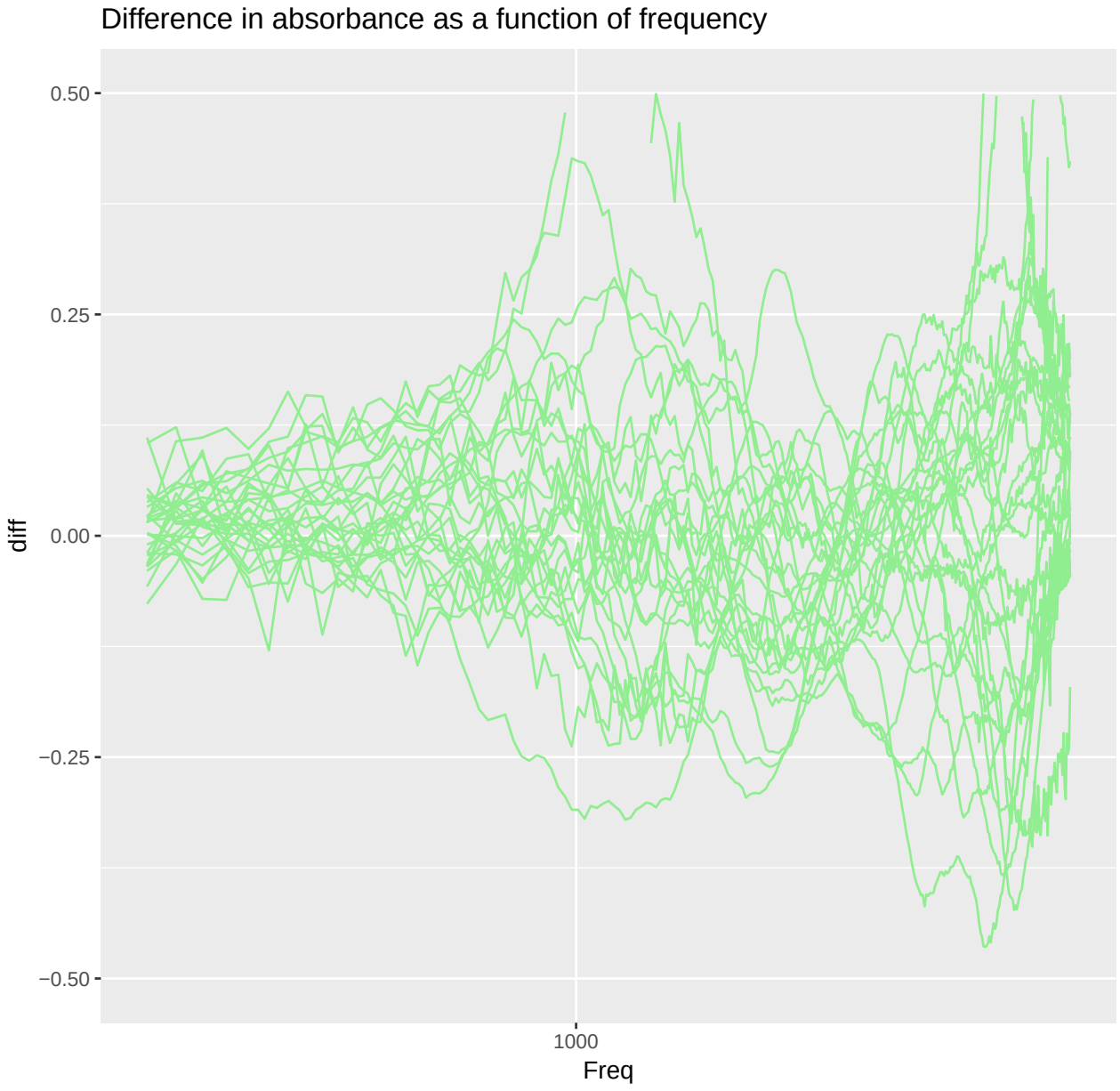



```
gf_line(Absorbance ~ Freq, color = ~ Sub_Number, group = ~ Sub_Number,
        data = Rosowski) %>%
  gf_facet_wrap(~ EarStatus) %>%
  gf_refine(scale_x_log10()) %>%
  gf_labs(title="Absorbance as a function of frequency",
       caption = "Data from Rosowski et al, 2012")
```


```
library(tidyr)
Rosowskiwide <- Rosowski %>%
  select(Sub_Number, EarStatus, Freq, Absorbance) %>%
  spread(EarStatus, Absorbance) %>%
  mutate(diff = Left - Right)
```

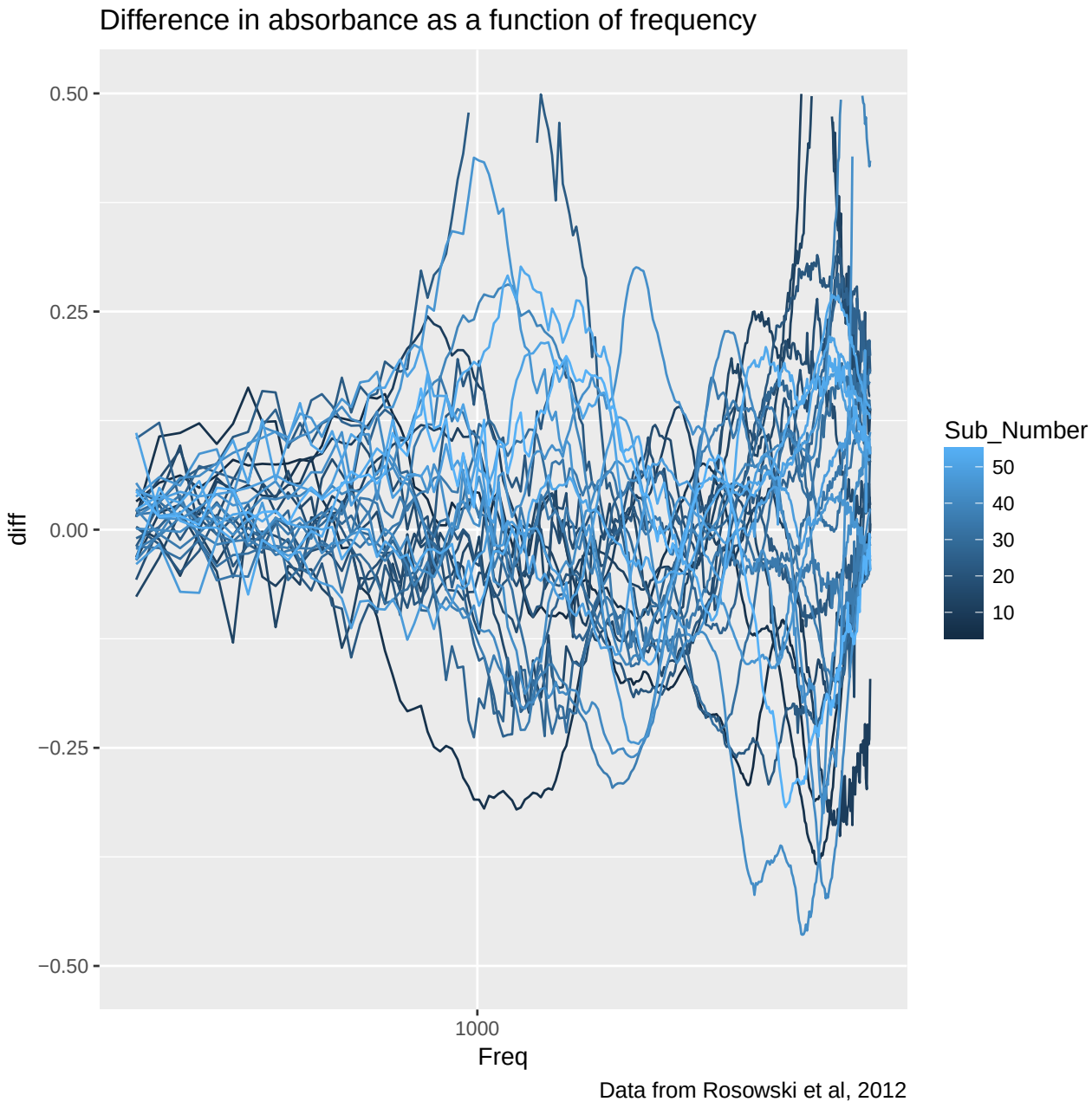
```
ggplot(Rosowskiwide, aes(x=Freq, y=diff, group=Sub_Number)) +
  geom_line(colour="lightgreen") +
  scale_x_log10() +
  ylim(-0.5, 0.5) +
  labs(title="Difference in absorbance as a function of frequency", caption = "Data from Rosowski et al
```

```
## Warning: Removed 20 rows containing missing values (geom_path).
```




```
gf_line(diff ~ Freq, group = ~ Sub_Number, color = ~ Sub_Number, data = Rosowskiwide) %>%
  gf_refine(scale_x_log10()) %>%
  gf_lims(y = c(-0.5, 0.5)) %>%
  gf_labs(title="Difference in absorbance as a function of frequency",
          caption = "Data from Rosowski et al, 2012")
```

```
## Warning: Removed 20 rows containing missing values (geom_path).
```


Difference in absorbance as a function of frequency
0.50
0.25
0.00
−0.25
−0.50
diff
1000
Freq
Sub_Number
50
40
30
20
10
Data from Rosowski et al, 2012