Table of Contents

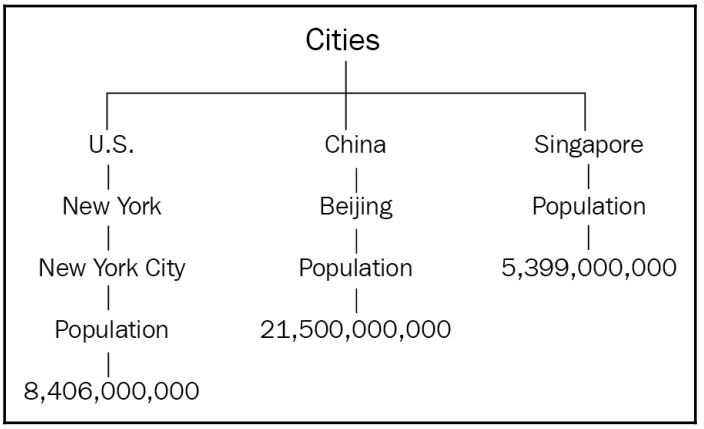
Chapter 1: Graphics bundle 1

- **Chapter 01: Programming with Data** 1
- **Chapter 02: Introduction to Programming in Python** 2
- **Chapter 03: Reading, Exploring, and Modifying Data - Part I** 6
- **Chapter 04: Reading, Exploring, and Modifying Data - Part II** 15
- **Chapter 05: Manipulating Text Data - An Introduction to Regular Expressions** 19
- **Chapter 06: Cleaning Numerical Data - An Introduction to R and RStudio** 21
- **Chapter 07: Simplifying Data Manipulation with dplyr** 28
- **Chapter 08: Getting Data from the Web** 32
- **Chapter 09: Working with Large Datasets** 36

Index 38

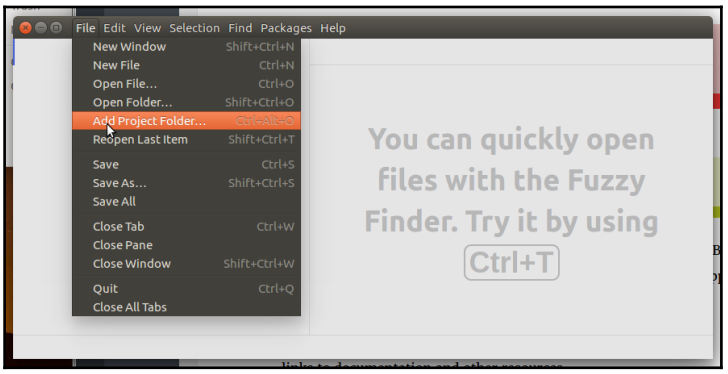
1
Graphics bundle

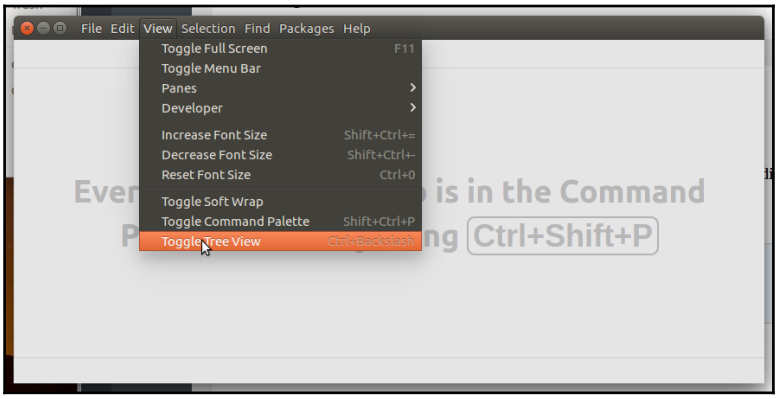
Chapter 01: Programming with Data

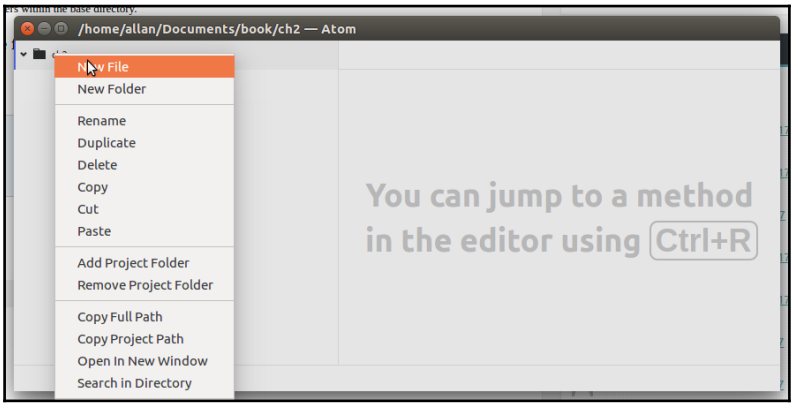


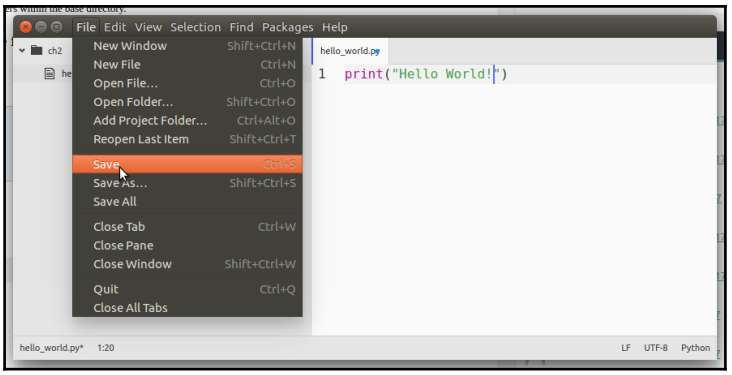

```
[ 8406000000 , 21500000000 , 5399000000 ]
```

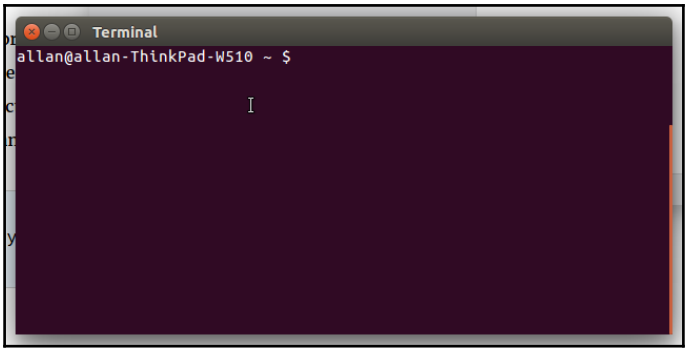
Chapter 02: Introduction to Programming in Python

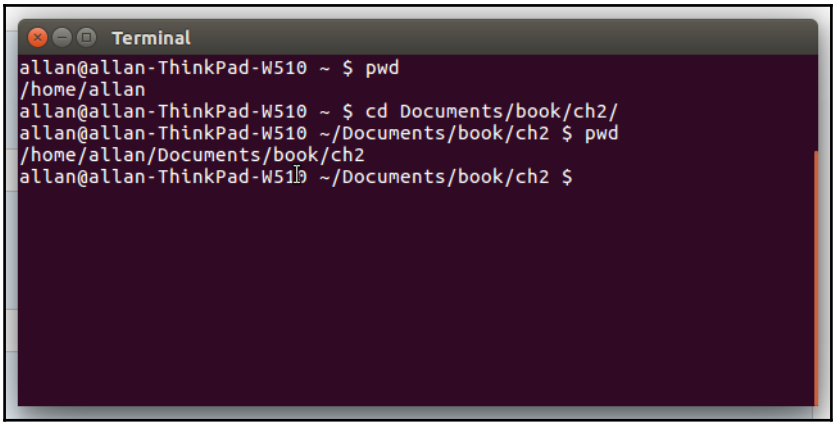


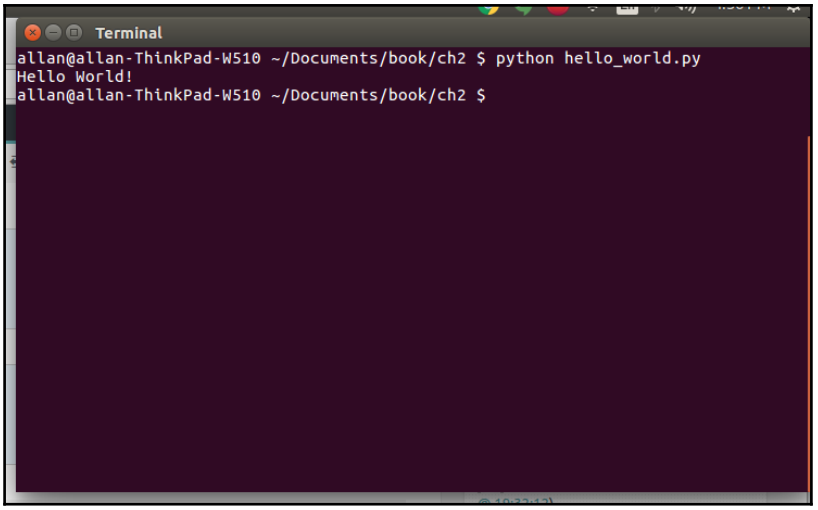


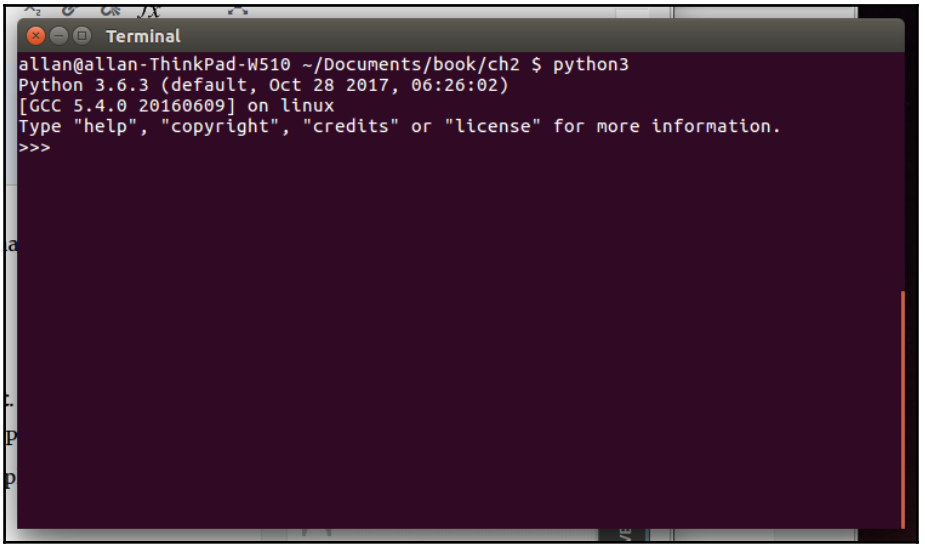


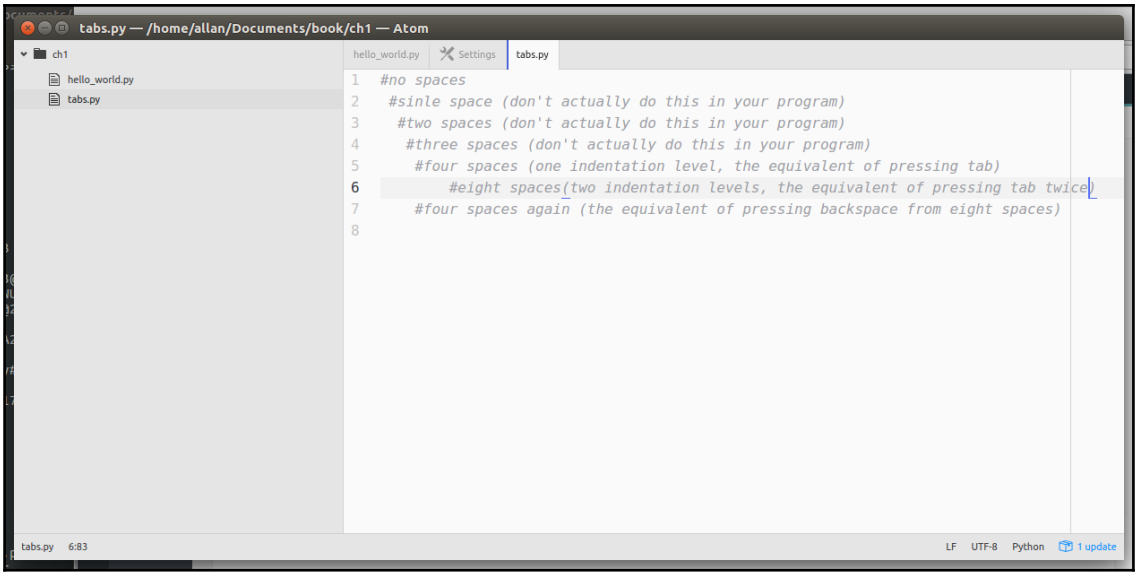

File Edit View Selection Find Packages Help
New Window Shift+Ctrl+N
New File Ctrl+N
Open File... Ctrl+O
Open Folder... Shift+Ctrl+O
Add Project Folder... Ctrl+Alt+O
Reopen Last Item Shift+Ctrl+T
Save Ctrl+S
Save As... Shift+Ctrl+S
Save All
Close Tab Ctrl+W
Close Pane
Close Window Shift+Ctrl+W
Quit Ctrl+Q
Close All Tabs
hello_world.py
1 print("Hello World!")
hello_world.py* 1:20
LF UTF-8 Python

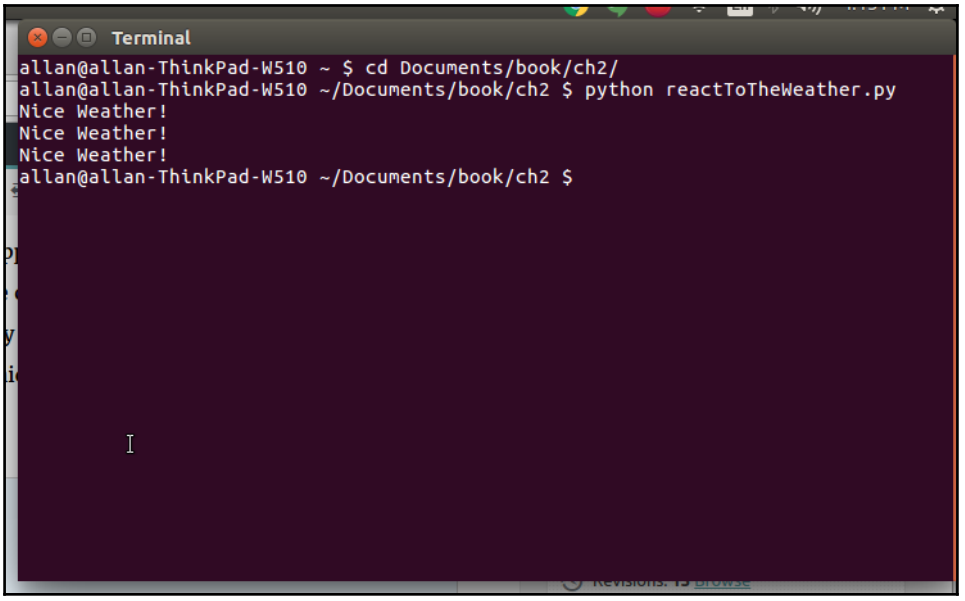

Terminal
allan@allan-ThinkPad-W510 ~ $


Terminal
allan@allan-ThinkPad-W510 ~ $ pwd
/home/allan
allan@allan-ThinkPad-W510 ~ $ cd Documents/book/ch2/
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $ pwd
/home/allan/Documents/book/ch2
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $


Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $ python hello_world.py
Hello World!
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $

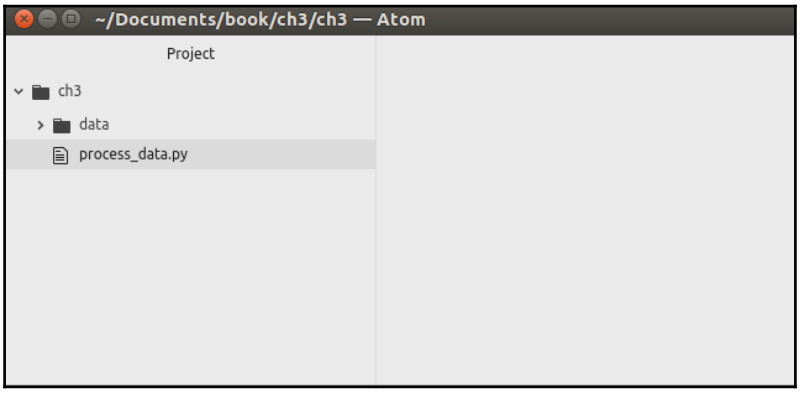

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $ python3
Python 3.6.3 (default, Oct 28 2017, 06:26:02)
[GCC 5.4.0 20160609] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>>

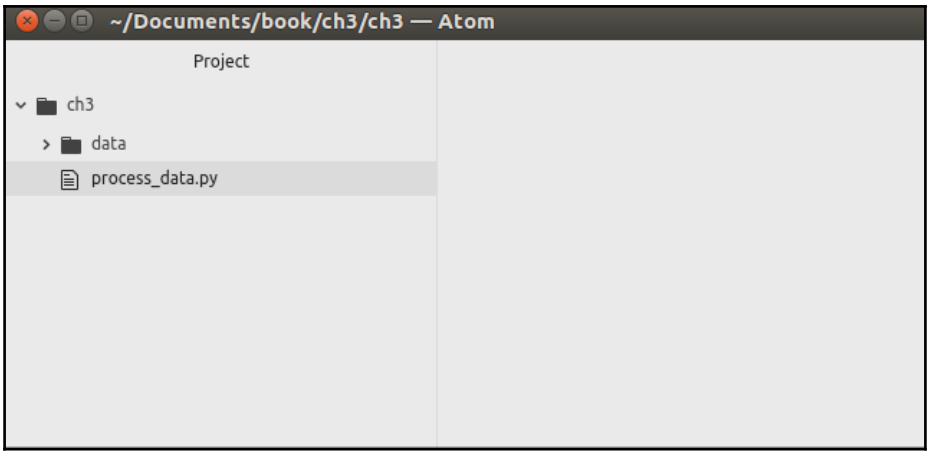

tabs.py — /home/allan/Documents/book/ch1 — Atom
ch1
hello_world.py
tabs.py
hello_world.py
Settings
tabs.py
1 #no spaces
2 #sinle space (don't actually do this in your program)
3 #two spaces (don't actually do this in your program)
4 #three spaces (don't actually do this in your program)
5 #four spaces (one indentation level, the equivalent of pressing tab)
6 #eight spaces(two indentation levels, the equivalent of pressing tab twice)
7 #four spaces again (the equivalent of pressing backspace from eight spaces)
8
tabs.py 6:83
LF UTF-8 Python 1 update

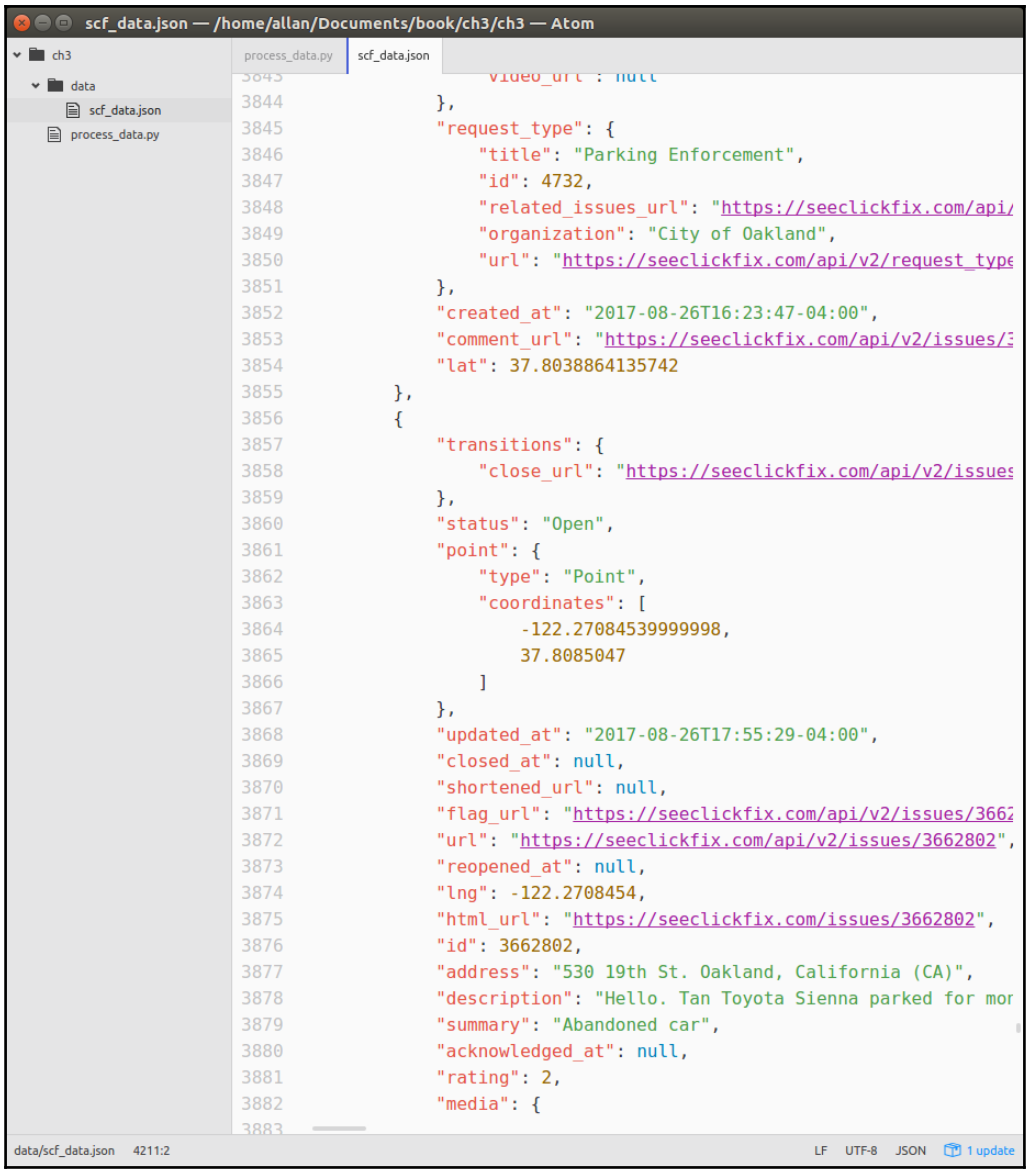

Terminal
allan@allan-ThinkPad-W510 ~ $ cd Documents/book/ch2/
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $ python reactToTheWeather.py
Nice Weather!
Nice Weather!
Nice Weather!
allan@allan-ThinkPad-W510 ~/Documents/book/ch2 $

Chapter 03: Reading, Exploring, and Modifying Data - Part I

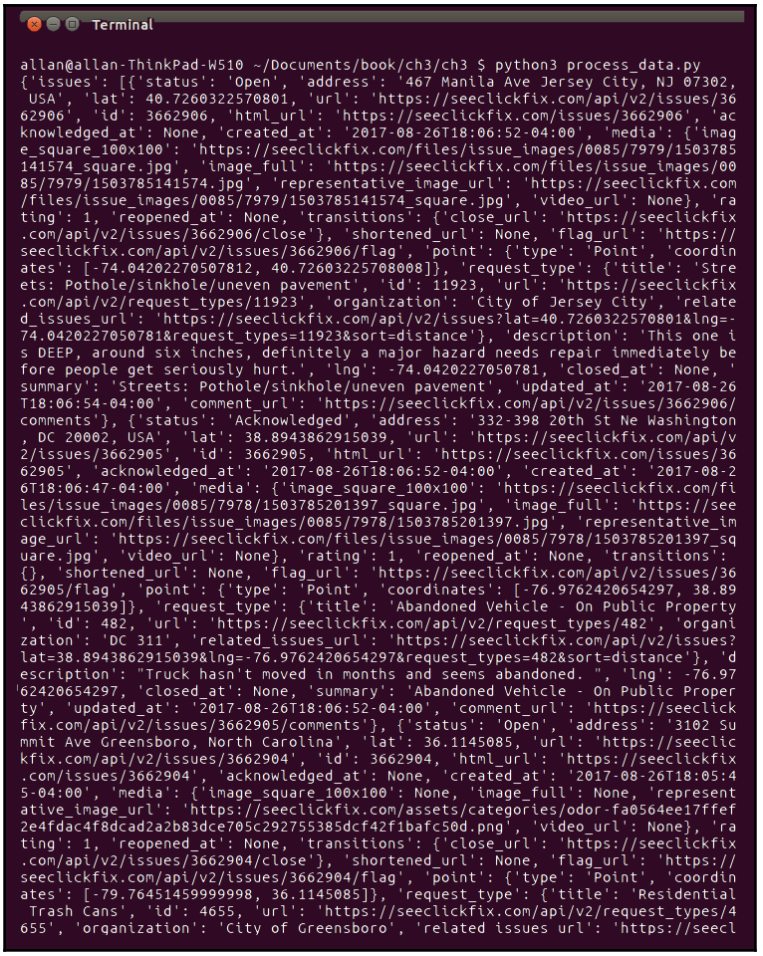

```
},
"updated_at": "2017-08-26
"closed_at": null,
"shortened_url": null,
"flag_url": "https://seec
"url": "https://seeclickf
"reopened_at": null,
"lng": -74.0420227050781,
"html_url": "https://seec
"id": 3662906,
"address": "467 Manila Av
"description": "This one
"summary": "Streets: Poth
"acknowledged_at": null,
"rating": 1,
"media": {
    "image_square_100x100
    "representative_image
    "image_full": "https:
    "video_url": null
},
```

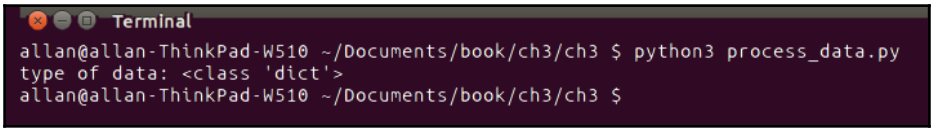



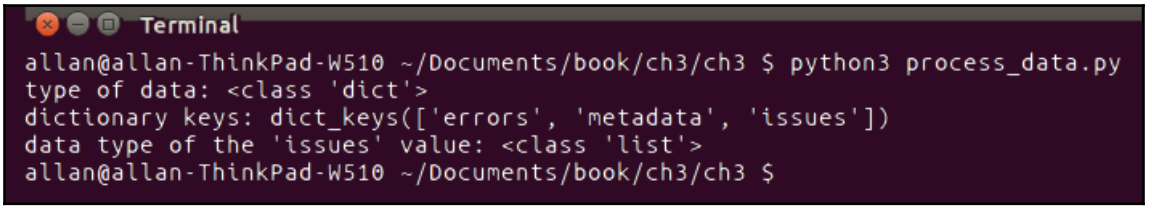

~/Documents/book/ch3/ch3 — Atom
Project
ch3
data
process_data.py

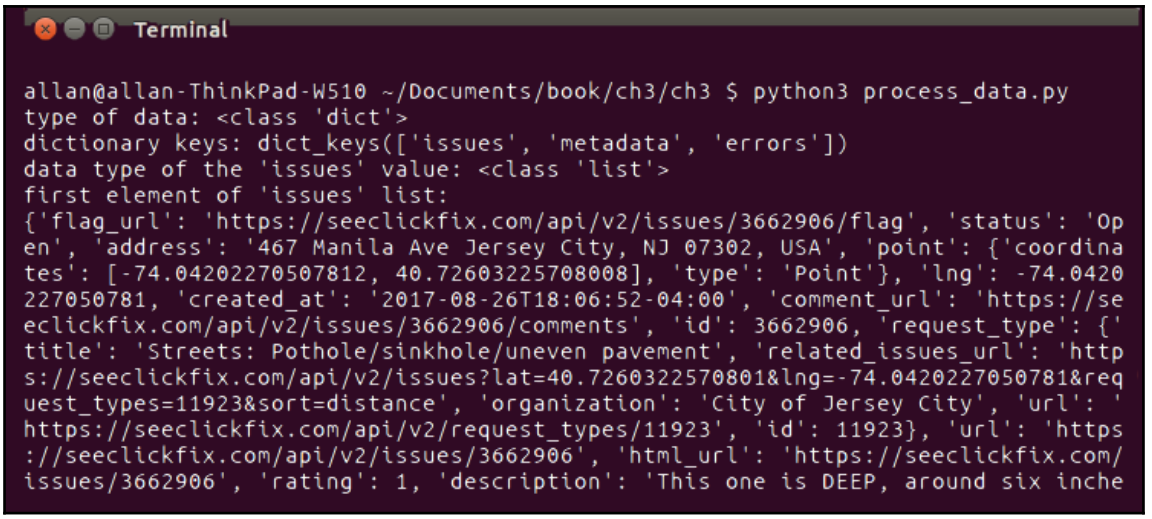

scf_data.json — /home/allan/Documents/book/ch3/ch3 — Atom
ch3
data
scf_data.json
process_data.py
process_data.py
scf_data.json
"video_url": null
},
"request_type": {
"title": "Parking Enforcement",
"id": 4732,
"related_issues_url": "https://seeclickfix.com/api/
"organization": "City of Oakland",
"url": "https://seeclickfix.com/api/v2/request_type
},
"created_at": "2017-08-26T16:23:47-04:00",
"comment_url": "https://seeclickfix.com/api/v2/issues/3
"lat": 37.8038864135742
},
{
"transitions": {
"close_url": "https://seeclickfix.com/api/v2/issues
},
"status": "Open",
"point": {
"type": "Point",
"coordinates": [
-122.27084539999998,
37.8085047
]
},
"updated_at": "2017-08-26T17:55:29-04:00",
"closed_at": null,
"shortened_url": null,
"flag_url": "https://seeclickfix.com/api/v2/issues/3662
"url": "https://seeclickfix.com/api/v2/issues/3662802",
"reopened_at": null,
"lng": -122.2708454,
"html_url": "https://seeclickfix.com/issues/3662802",
"id": 3662802,
"address": "530 19th St. Oakland, California (CA)",
"description": "Hello. Tan Toyota Sienna parked for mor
"summary": "Abandoned car",
"acknowledged_at": null,
"rating": 2,
"media": {
data/scf_data.json 4211:2
LF UTF-8 JSON 1 update

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
{
"errors": {},
"issues": [
{
"transitions": {
"close_url": "https://seeclickfix.com/api/v2/issues/3662906/close"
},
"status": "Open",
"point": {
"type": "Point",
"coordinates": [
-74.04202270507812,
40.72603225708008
]
},
"updated_at": "2017-08-26T18:06:54-04:00",
"closed_at": null,
"shortened_url": null,
"flag_url": "https://seeclickfix.com/api/v2/issues/3662906/flag",
"url": "https://seeclickfix.com/api/v2/issues/3662906",
"reopened_at": null,
"lng": -74.0420227050781,
"html_url": "https://seeclickfix.com/issues/3662906",
"id": 3662906,
"address": "467 Manila Ave Jersey City, NJ 07302, USA",
"description": "This one is DEEP, around six inches, definitely a major hazard needs repair immediately before people get seriously hurt.",
"summary": "Streets: Pothole/sinkhole/uneven pavement",
"acknowledged_at": null,
"rating": 1,
"media": {
"image_square_100x100": "https://seeclickfix.com/files/issue_images/0085/7979/1503785141574_square.jpg",
"representative_image_url": "https://seeclickfix.com/files/issue_images/0085/7979/1503785141574_square.jpg",
"image_full": "https://seeclickfix.com/files/issue_images/0085/7979/1503785141574.jpg",
"video_url": null
},
"request_type": {
"title": "Streets: Pothole/sinkhole/uneven pavement",
"id": 11923,
"related_issues_url": "https://seeclickfix.com/api/v2/issues?lat=40.72603225708001&lng=-74.04202270507781&request_types=11923&sort=distance",
"organization": "City of Jersey City",
"url": "https://seeclickfix.com/api/v2/request_types/11923"
},
"created_at": "2017-08-26T18:06:52-04:00",
"comment_url": "https://seeclickfix.com/api/v2/issues/3662906/commen


Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
{'issues': [{'status': 'Open', 'address': '467 Manila Ave Jersey City, NJ 07302,
USA', 'lat': 40.7260322570801, 'url': 'https://seeclickfix.com/api/v2/issues/36
62906', 'id': 3662906, 'html_url': 'https://seeclickfix.com/issues/3662906', 'ac
knowledged_at': None, 'created_at': '2017-08-26T18:06:52-04:00', 'media': {'imag
e_square_100x100': 'https://seeclickfix.com/files/issue_images/0085/7979/1503785
141574_square.jpg', 'image_full': 'https://seeclickfix.com/files/issue_images/00
85/7979/1503785141574.jpg', 'representative_image_url': 'https://seeclickfix.com
/files/issue_images/0085/7979/1503785141574_square.jpg', 'video_url': None}, 'ra
ting': 1, 'reopened_at': None, 'transitions': {'close_url': 'https://seeclickfix
.com/api/v2/issues/3662906/close'}, 'shortened_url': None, 'flag_url': 'https://
seeclickfix.com/api/v2/issues/3662906/flag', 'point': {'type': 'Point', 'coordin
ates': [-74.04202270507812, 40.72603225708008]}, 'request_type': {'title': 'Stre
ets: Pothole/sinkhole/uneven pavement', 'id': 11923, 'url': 'https://seeclickfix
.com/api/v2/request_types/11923', 'organization': 'City of Jersey City', 'relate
d_issues_url': 'https://seeclickfix.com/api/v2/issues?lat=40.7260322570801&lng=-
74.0420227050781&request_types=11923&sort=distance'}, 'description': 'This one i
s DEEP, around six inches, definitely a major hazard needs repair immediately be
fore people get seriously hurt.', 'lng': -74.0420227050781, 'closed_at': None, '
summary': 'Streets: Pothole/sinkhole/uneven pavement', 'updated_at': '2017-08-26
T18:06:54-04:00', 'comment_url': 'https://seeclickfix.com/api/v2/issues/3662906/
comments'}, {'status': 'Acknowledged', 'address': '332-398 20th St Ne Washington
, DC 20002, USA', 'lat': 38.8943862915039, 'url': 'https://seeclickfix.com/api/v
2/issues/3662905', 'id': 3662905, 'html_url': 'https://seeclickfix.com/issues/36
62905', 'acknowledged_at': '2017-08-26T18:06:52-04:00', 'created_at': '2017-08-2
6T18:06:47-04:00', 'media': {'image_square_100x100': 'https://seeclickfix.com/fi
les/issue_images/0085/7978/1503785201397_square.jpg', 'image_full': 'https://see
clickfix.com/files/issue_images/0085/7978/1503785201397.jpg', 'representative_im
age_url': 'https://seeclickfix.com/files/issue_images/0085/7978/1503785201397_sq
uare.jpg', 'video_url': None}, 'rating': 1, 'reopened_at': None, 'transitions':
{}, 'shortened_url': None, 'flag_url': 'https://seeclickfix.com/api/v2/issues/36
62905/flag', 'point': {'type': 'Point', 'coordinates': [-76.9762420654297, 38.89
43862915039]}, 'request_type': {'title': 'Abandoned Vehicle - On Public Property
', 'id': 482, 'url': 'https://seeclickfix.com/api/v2/request_types/482', 'organi
zation': 'DC 311', 'related_issues_url': 'https://seeclickfix.com/api/v2/issues?
lat=38.8943862915039&lng=-76.9762420654297&request_types=482&sort=distance'}, 'd
escription': "Truck hasn't moved in months and seems abandoned. ", 'lng': -76.97
62420654297, 'closed_at': None, 'summary': 'Abandoned Vehicle - On Public Proper
ty', 'updated_at': '2017-08-26T18:06:52-04:00', 'comment_url': 'https://seeclick
fix.com/api/v2/issues/3662905/comments'}, {'status': 'Open', 'address': '3102 Su
mmit Ave Greensboro, North Carolina', 'lat': 36.1145085, 'url': 'https://seeclic
kfix.com/api/v2/issues/3662904', 'id': 3662904, 'html_url': 'https://seeclickfix
.com/issues/3662904', 'acknowledged_at': None, 'created_at': '2017-08-26T18:05:4
5-04:00', 'media': {'image_square_100x100': None, 'image_full': None, 'represent
ative_image_url': 'https://seeclickfix.com/assets/categories/odor-fa0564ee17ffef
2e4fdac4f8dcad2a2b83dce705c292755385dcf42f1bafc50d.png', 'video_url': None}, 'ra
ting': 1, 'reopened_at': None, 'transitions': {'close_url': 'https://seeclickfix
.com/api/v2/issues/3662904/close'}, 'shortened_url': None, 'flag_url': 'https://
seeclickfix.com/api/v2/issues/3662904/flag', 'point': {'type': 'Point', 'coordin
ates': [-79.76451459999998, 36.1145085]}, 'request_type': {'title': 'Residential
Trash Cans', 'id': 4655, 'url': 'https://seeclickfix.com/api/v2/request_types/4
655', 'organization': 'City of Greensboro', 'related_issues_url': 'https://seecl

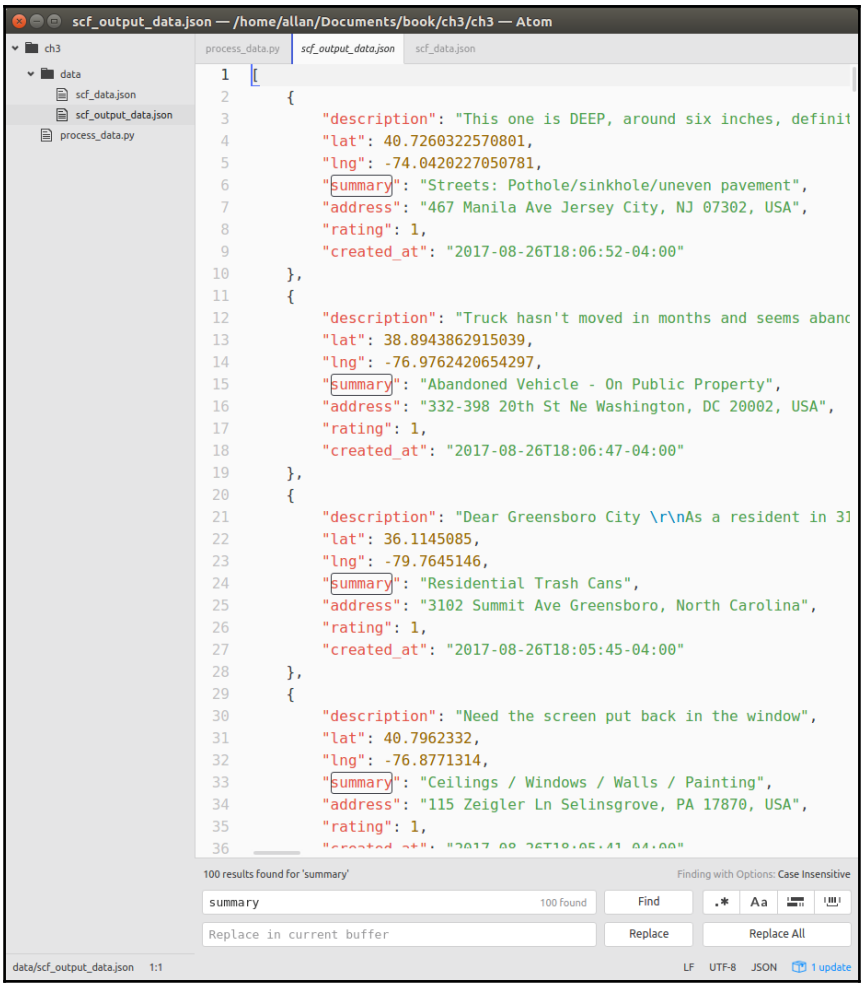

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
type of data: <class 'dict'>
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $

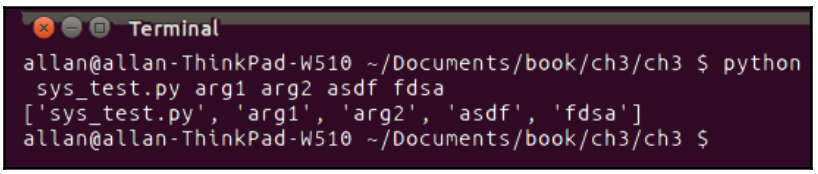

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
type of data: <class 'dict'>
dictionary keys: dict_keys(['errors', 'metadata', 'issues'])
data type of the 'issues' value: <class 'list'>
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $


Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
type of data: <class 'dict'>
dictionary keys: dict_keys(['issues', 'metadata', 'errors'])
data type of the 'issues' value: <class 'list'>
first element of 'issues' list:
{'flag_url': 'https://seeclickfix.com/api/v2/issues/3662906/flag', 'status': 'Op
en', 'address': '467 Manila Ave Jersey City, NJ 07302, USA', 'point': {'coordina
tes': [-74.04202270507812, 40.72603225708008], 'type': 'Point'}, 'lng': -74.0420
227050781, 'created_at': '2017-08-26T18:06:52-04:00', 'comment_url': 'https://se
eclickfix.com/api/v2/issues/3662906/comments', 'id': 3662906, 'request_type': {'
title': 'Streets: Pothole/sinkhole/uneven pavement', 'related_issues_url': 'http
s://seeclickfix.com/api/v2/issues?lat=40.7260322570801&lng=-74.0420227050781&req
uest_types=11923&sort=distance', 'organization': 'City of Jersey City', 'url': '
https://seeclickfix.com/api/v2/request_types/11923', 'id': 11923}, 'url': 'https
://seeclickfix.com/api/v2/issues/3662906', 'html_url': 'https://seeclickfix.com/
issues/3662906', 'rating': 1, 'description': 'This one is DEEP, around six inche

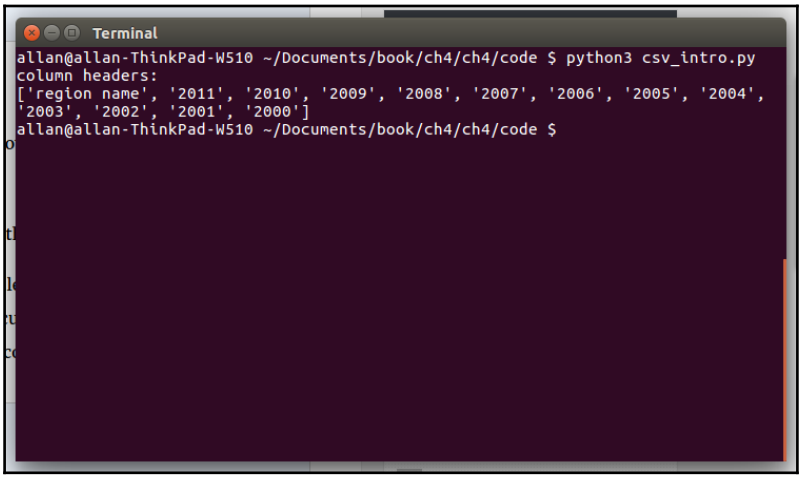
```
Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
first data entry:
{   'acknowledged_at': None,
    'address': '467 Manila Ave Jersey City, NJ 07302, USA',
    'closed_at': None,
    'comment_url': 'https://seeclickfix.com/api/v2/issues/3662906/comments',
    'created_at': '2017-08-26T18:06:52-04:00',
    'description': 'This one is DEEP, around six inches, definitely a major '
                   'hazard needs repair immediately before people get '
                   'seriously hurt.',
    'flag_url': 'https://seeclickfix.com/api/v2/issues/3662906/flag',
    'html_url': 'https://seeclickfix.com/issues/3662906',
    'id': 3662906,
    'lat': 40.7260322570801,
    'lng': -74.0420227050781,
    'media': {   'image_full': 'https://seeclickfix.com/files/issue_images/0085/
7979/1503785141574.jpg',
                 'image_square_100x100': 'https://seeclickfix.com/files/issue_im
ages/0085/7979/1503785141574_square.jpg',
                 'representative_image_url': 'https://seeclickfix.com/files/issu
e_images/0085/7979/1503785141574_square.jpg',
                 'video_url': None},
    'point': {   'coordinates': [-74.04202270507812, 40.72603225708008],
                 'type': 'Point'},
    'rating': 1,
    'reopened_at': None,
    'request_type': {   'id': 11923,
                        'organization': 'City of Jersey City',
                        'related_issues_url': 'https://seeclickfix.com/api/v2/is
sues?lat=40.7260322570801&lng=-74.0420227050781&request_types=11923&sort=distanc
e',
                        'title': 'Streets: Pothole/sinkhole/uneven pavement',
                        'url': 'https://seeclickfix.com/api/v2/request_types/119
23'},
    'shortened_url': None,
    'status': 'Open',
    'summary': 'Streets: Pothole/sinkhole/uneven pavement',
    'transitions': {   'close_url': 'https://seeclickfix.com/api/v2/issues/36629
06/close'},
    'updated_at': '2017-08-26T18:06:54-04:00',
    'url': 'https://seeclickfix.com/api/v2/issues/3662906'}
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $
```

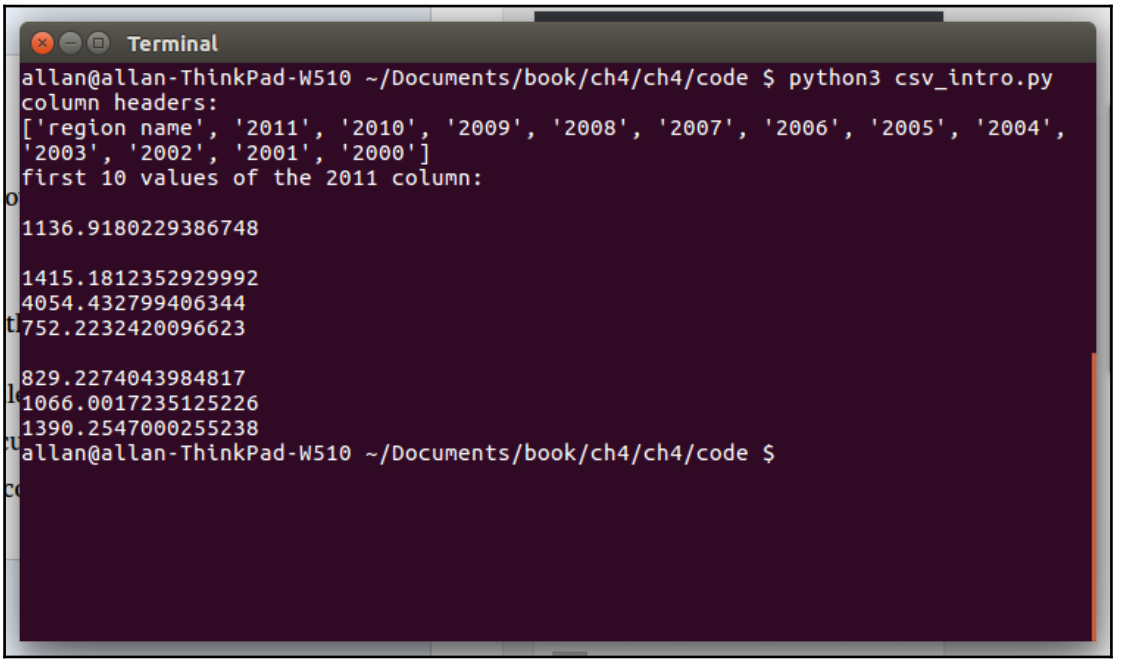
```
Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python3 process_data.py
{'address': '467 Manila Ave Jersey City, NJ 07302, USA', 'created_at': '2017-08-26T18:06:52-04:00', 'description': 'This one
 is DEEP, around six inches, definitely a major hazard needs repair immediately before people get seriously hurt.', 'lng': -
74.0420227050781, 'rating': 1, 'lat': 40.7260322570801}
{'address': '332-398 20th St Ne Washington, DC 20002, USA', 'created_at': '2017-08-26T18:06:47-04:00', 'description': "Truck
 hasn't moved in months and seems abandoned. ", 'lng': -76.9762420654297, 'rating': 1, 'lat': 38.8943862915039}
{'address': '3102 Summit Ave Greensboro, North Carolina', 'created_at': '2017-08-26T18:05:45-04:00', 'description': "Dear Gr
eensboro City \r\nAs a resident in 3100-3102-3104 Summit avenue \r\nthe Trash collect cars don't come as schedule on Thursda
y\r\nThe street is full by trash\r\nand the insects spread everywhere specially in this hot weather in add to the bad smell\
r\nthat why we ask you to send the trash collector cars immediately and we ask you to provide another trash can since three
are not enough for about 200 resident in these area\r\nThanks \r\n", 'lng': -79.7645146, 'rating': 1, 'lat': 36.1145085}
{'address': '115 Zeigler Ln Selinsgrove, PA 17870, USA', 'created_at': '2017-08-26T18:05:41-04:00', 'description': 'Need the
 screen put back in the window', 'lng': -76.8771314, 'rating': 1, 'lat': 40.7962332}
{'address': '921 E Clovefield St Gilbert, Arizona', 'created_at': '2017-08-26T18:04:55-04:00', 'description': 'Think we have
 a water leak.  Had a Red Push tree that died quickly.  Bought new tree to replace however the hole keeps filling up with wa
ter and we have turned off our water source.', 'lng': -111.7701027, 'rating': 1, 'lat': 33.2571791}
{'address': '305 Longwood Ave  Huntsville, Alabama', 'created_at': '2017-08-26T18:04:26-04:00', 'description': ' Old Utility
 pole has broken and fell a portion is propped up and tied with wire', 'lng': -86.58382, 'rating': 1, 'lat': 34.7187132}
{'address': '900 70th Street North Saint Petersburg, Florida', 'created_at': '2017-08-26T18:04:01-04:00', 'description': 'La
rge amount of furniture abadoned by dumpster. Apartment residents cannot access dumpster.', 'lng': -82.7364878, 'rating': 1,
 'lat': 27.7814312}
{'address': '3212 Hyde St', 'created_at': '2017-08-26T18:03:59-04:00', 'description': 'Old chrysler? Car been here forever.
 Grass is growing under it.  Never gets moved not even for street cleaning.  Yellow and gray? Lp is 882 UAF.  ', 'lng': -122
.2204001, 'rating': 1, 'lat': 37.7884642069597}
{'address': '332 20th St Ne Washington, DC 20002, USA', 'created_at': '2017-08-26T18:03:51-04:00', 'description': '', 'lng':
-76.9763385742215, 'rating': 1, 'lat': 38.8943513004548}
{'address': '5745 Shasta Pines Valdosta, GA 31601, USA', 'created_at': '2017-08-26T18:02:31-04:00', 'description': "There ar
e several lots that are not maintained on a regular basis. Most are owned by one individual who owns several homes and lots.
 I don't want to say his name. Please have him to keep the lots cleaned. Thanks", 'lng': -83.2324265, 'rating': 1, 'lat': 30
.8166589}
{'address': '1200-1298 W 13th St Houston, TX 77008, USA', 'created_at': '2017-08-26T17:59:33-04:00', 'description': 'Gushing
 water for 3 days in front a driveway to Love Elementary ', 'lng': -95.4092944167996, 'rating': 1, 'lat': 29.7941687760981}
{'address': '4800 West Washington St S Broken Arrow, Oklahoma', 'created_at': '2017-08-26T17:56:08-04:00', 'description': 'M
any of us have been complaining regarding the fountains not working in Jackson park.after several complains including phone
calls the issue has not been resolved.We were told that parts had been ordered and it would take 6 weeks .I specifically hav
e been complaining over  a year now.You do understand that still water gives rise to algae and can harbor toxins.\r\nThere a
re so many mosquitoes in our backyard.Its impossible to enjoy summer especially for senior citizens\r\n Please address this
issue ASAP', 'lng': -95.8494697, 'rating': 1, 'lat': 36.0319415}
{'address': '1209-1213 Walton Dr College Station, TX 77840, USA', 'created_at': '2017-08-26T17:55:30-04:00', 'description':
'Cars parking in yard', 'lng': -96.319630990152, 'rating': 2, 'lat': 30.624716313137}
{'address': 'Calvert And Tunlaw Washington, District of Columbia', 'created_at': '2017-08-26T17:55:09-04:00', 'description':
 'faded crosswalks at Calvert and Tunlaw NW. please repaint. thanks.', 'lng': -77.0747085, 'rating': 1, 'lat': 38.9225454}
{'address': '39 And Davis Nw Washington, District of Columbia', 'created_at': '2017-08-26T17:54:09-04:00', 'description': 'f
aded crosswalks at 39 and Davis NW. please repaint. thanks.', 'lng': -77.0772463, 'rating': 1, 'lat': 38.923624}
{'address': '300 East Brookland Park Boulevard Richmond, Virginia', 'created_at': '2017-08-26T17:52:20-04:00', 'description'
: 'Grass is taller than 4 foot fence. In yard and between sidewalk and curb.', 'lng': -77.429018, 'rating': 1, 'lat': 37.571
547}
```

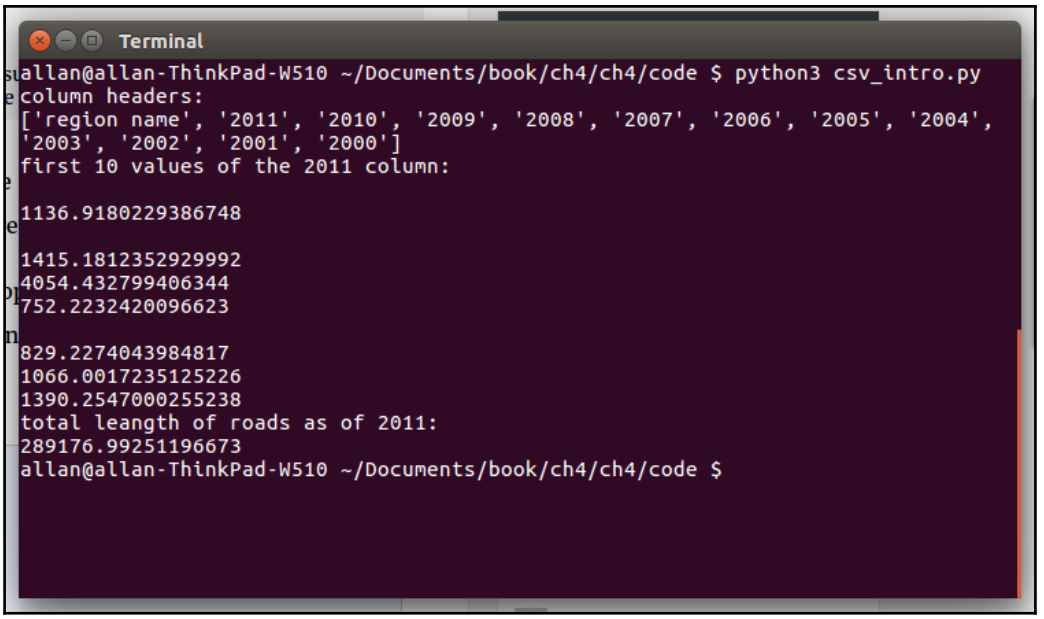

scf_output_data.json — /home/allan/Documents/book/ch3/ch3 — Atom
ch3
data
scf_data.json
scf_output_data.json
process_data.py
process_data.py
scf_output_data.json
scf_data.json
[
{
"description": "This one is DEEP, around six inches, definit
"lat": 40.7260322570801,
"lng": -74.0420227050781,
"summary": "Streets: Pothole/sinkhole/uneven pavement",
"address": "467 Manila Ave Jersey City, NJ 07302, USA",
"rating": 1,
"created_at": "2017-08-26T18:06:52-04:00"
},
{
"description": "Truck hasn't moved in months and seems abanc
"lat": 38.8943862915039,
"lng": -76.9762420654297,
"summary": "Abandoned Vehicle - On Public Property",
"address": "332-398 20th St Ne Washington, DC 20002, USA",
"rating": 1,
"created_at": "2017-08-26T18:06:47-04:00"
},
{
"description": "Dear Greensboro City \r\nAs a resident in 3
"lat": 36.1145085,
"lng": -79.7645146,
"summary": "Residential Trash Cans",
"address": "3102 Summit Ave Greensboro, North Carolina",
"rating": 1,
"created_at": "2017-08-26T18:05:45-04:00"
},
{
"description": "Need the screen put back in the window",
"lat": 40.7962332,
"lng": -76.8771314,
"summary": "Ceilings / Windows / Walls / Painting",
"address": "115 Zeigler Ln Selinsgrove, PA 17870, USA",
"rating": 1,
100 results found for 'summary'
Finding with Options: Case Insensitive
summary
100 found
Find
Aa
Replace in current buffer
Replace
Replace All
data/scf_output_data.json 1:1
LF UTF-8 JSON 1 update

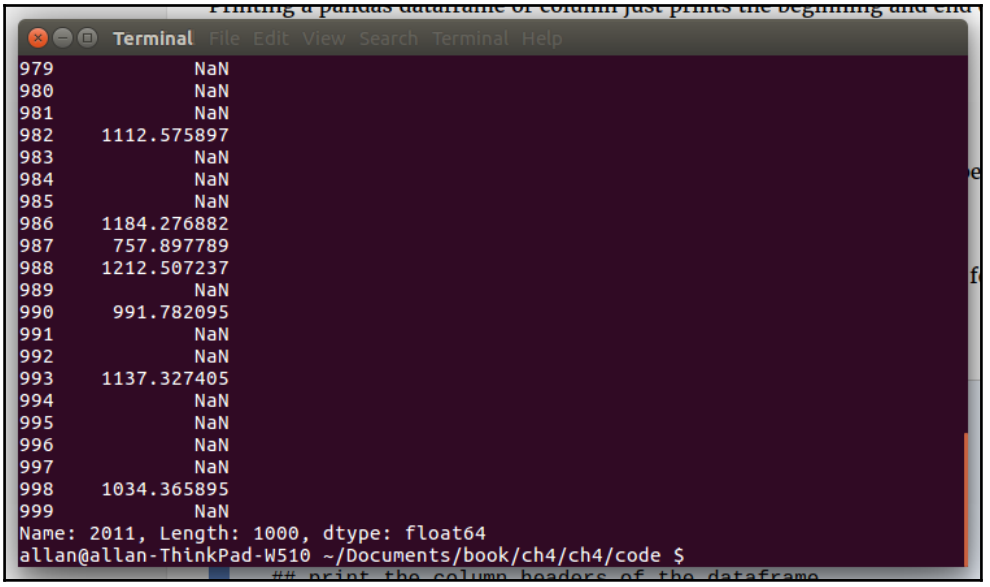

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $ python
sys_test.py arg1 arg2 asdf fdsa
['sys_test.py', 'arg1', 'arg2', 'asdf', 'fdsa']
allan@allan-ThinkPad-W510 ~/Documents/book/ch3/ch3 $

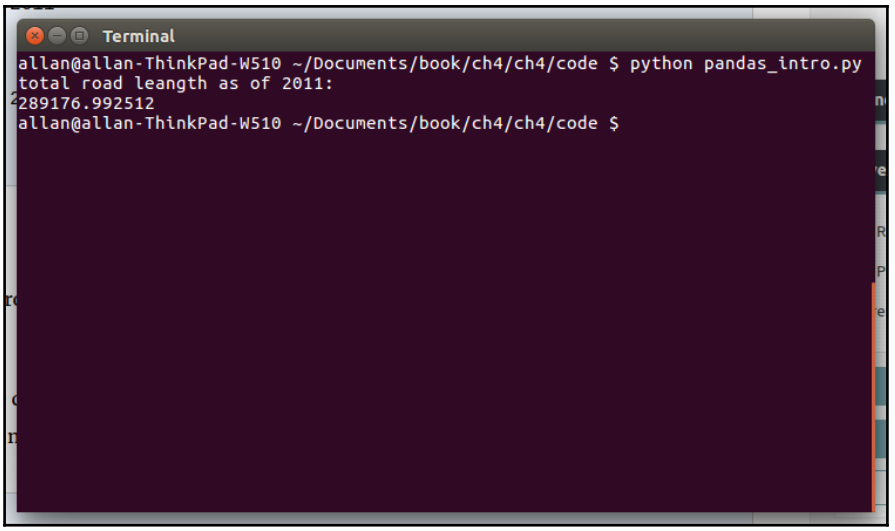
Chapter 04: Reading, Exploring, and Modifying Data - Part II

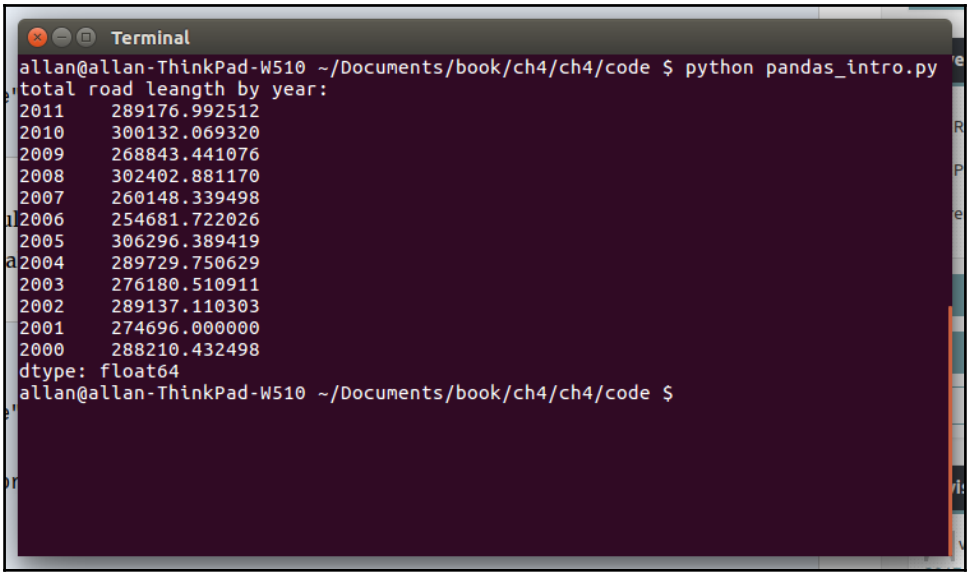


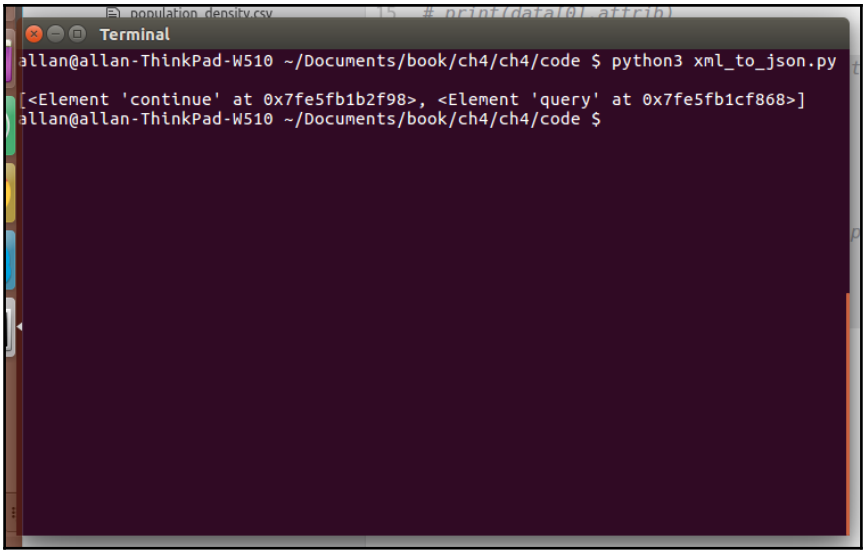


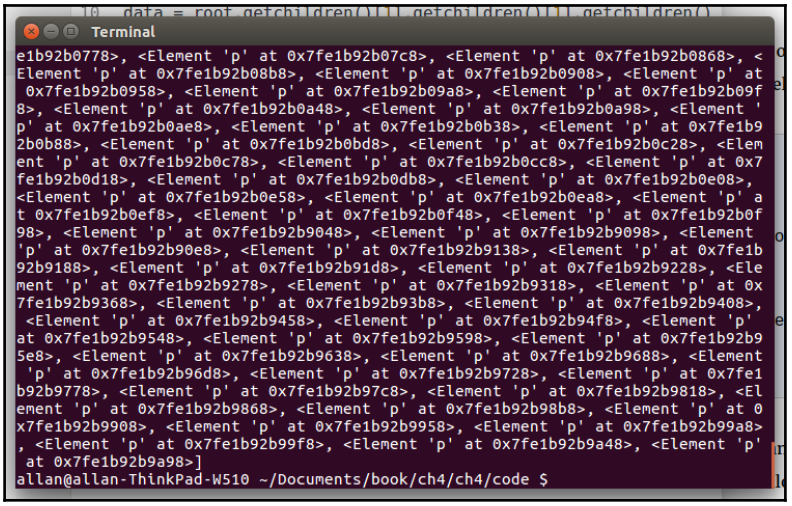

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $ python3 csv_intro.py
column headers:
['region name', '2011', '2010', '2009', '2008', '2007', '2006', '2005', '2004',
'2003', '2002', '2001', '2000']
first 10 values of the 2011 column:
1136.9180229386748
1415.1812352929992
4054.432799406344
752.2232420096623
829.2274043984817
1066.0017235125226
1390.2547000255238
total leangth of roads as of 2011:
289176.99251196673
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $


Terminal File Edit View Search Terminal Help
979 NaN
980 NaN
981 NaN
982 1112.575897
983 NaN
984 NaN
985 NaN
986 1184.276882
987 757.897789
988 1212.507237
989 NaN
990 991.782095
991 NaN
992 NaN
993 1137.327405
994 NaN
995 NaN
996 NaN
997 NaN
998 1034.365895
999 NaN
Name: 2011, Length: 1000, dtype: float64
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $

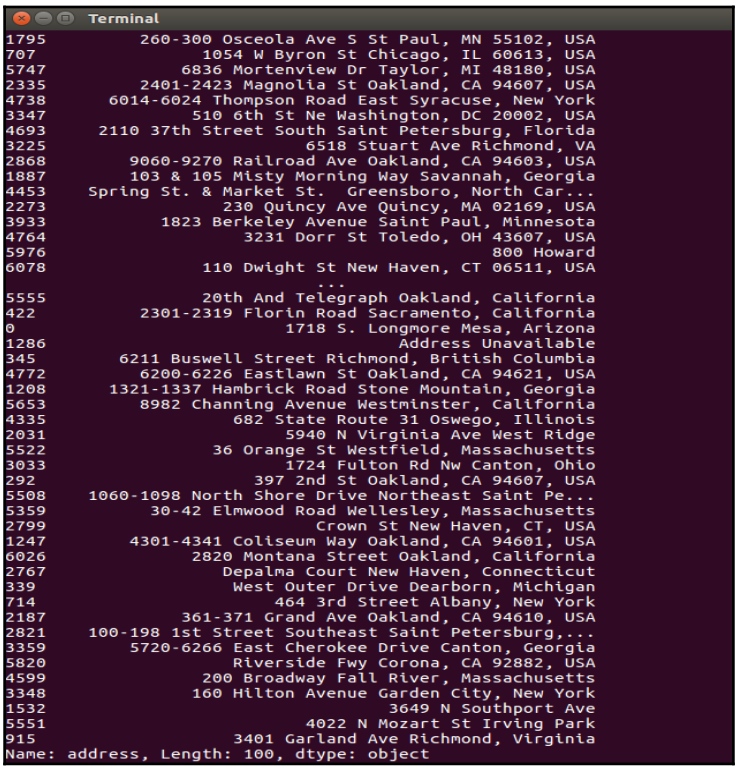

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $ python pandas_intro.py
total road leangth as of 2011:
289176.992512
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $

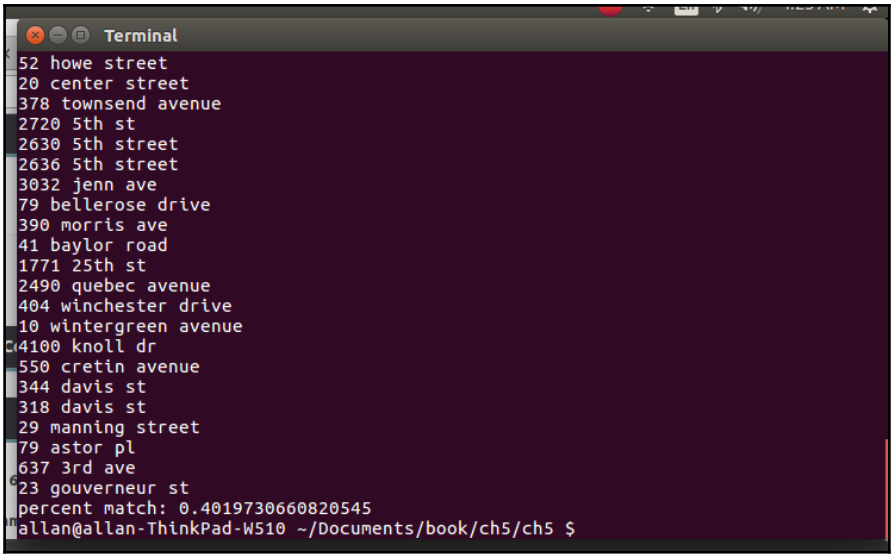

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $ python pandas_intro.py
total road leangth by year:
2011 289176.992512
2010 300132.069320
2009 268843.441076
2008 302402.881170
2007 260148.339498
2006 254681.722026
2005 306296.389419
2004 289729.750629
2003 276180.510911
2002 289137.110303
2001 274696.000000
2000 288210.432498
dtype: float64
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $

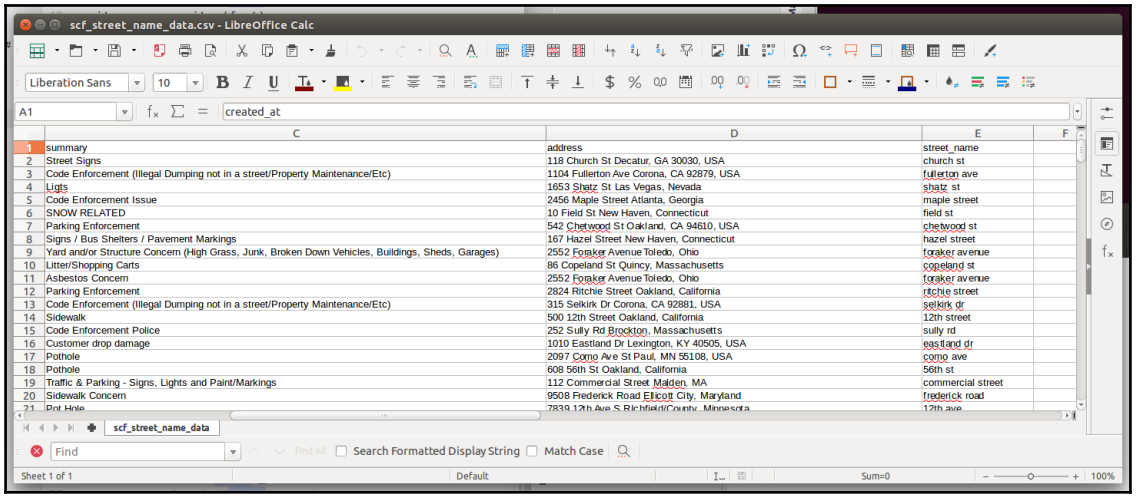

Terminal
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $ python3 xml_to_json.py
[<Element 'continue' at 0x7fe5fb1b2f98>, <Element 'query' at 0x7fe5fb1cf868>]
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $


Terminal
e1b92b0778>, <Element 'p' at 0x7fe1b92b07c8>, <Element 'p' at 0x7fe1b92b0868>, <
Element 'p' at 0x7fe1b92b08b8>, <Element 'p' at 0x7fe1b92b0908>, <Element 'p' at
0x7fe1b92b0958>, <Element 'p' at 0x7fe1b92b09a8>, <Element 'p' at 0x7fe1b92b09f
8>, <Element 'p' at 0x7fe1b92b0a48>, <Element 'p' at 0x7fe1b92b0a98>, <Element '
p' at 0x7fe1b92b0ae8>, <Element 'p' at 0x7fe1b92b0b38>, <Element 'p' at 0x7fe1b9
2b0b88>, <Element 'p' at 0x7fe1b92b0bd8>, <Element 'p' at 0x7fe1b92b0c28>, <Elem
ent 'p' at 0x7fe1b92b0c78>, <Element 'p' at 0x7fe1b92b0cc8>, <Element 'p' at 0x7
fe1b92b0d18>, <Element 'p' at 0x7fe1b92b0db8>, <Element 'p' at 0x7fe1b92b0e08>,
<Element 'p' at 0x7fe1b92b0e58>, <Element 'p' at 0x7fe1b92b0ea8>, <Element 'p' a
t 0x7fe1b92b0ef8>, <Element 'p' at 0x7fe1b92b0f48>, <Element 'p' at 0x7fe1b92b0f
98>, <Element 'p' at 0x7fe1b92b9048>, <Element 'p' at 0x7fe1b92b9098>, <Element
'p' at 0x7fe1b92b90e8>, <Element 'p' at 0x7fe1b92b9138>, <Element 'p' at 0x7fe1b
92b9188>, <Element 'p' at 0x7fe1b92b91d8>, <Element 'p' at 0x7fe1b92b9228>, <Ele
ment 'p' at 0x7fe1b92b9278>, <Element 'p' at 0x7fe1b92b9318>, <Element 'p' at 0x
7fe1b92b9368>, <Element 'p' at 0x7fe1b92b93b8>, <Element 'p' at 0x7fe1b92b9408>,
<Element 'p' at 0x7fe1b92b9458>, <Element 'p' at 0x7fe1b92b94f8>, <Element 'p'
at 0x7fe1b92b9548>, <Element 'p' at 0x7fe1b92b9598>, <Element 'p' at 0x7fe1b92b9
5e8>, <Element 'p' at 0x7fe1b92b9638>, <Element 'p' at 0x7fe1b92b9688>, <Element
'p' at 0x7fe1b92b96d8>, <Element 'p' at 0x7fe1b92b9728>, <Element 'p' at 0x7fe1
b92b9778>, <Element 'p' at 0x7fe1b92b97c8>, <Element 'p' at 0x7fe1b92b9818>, <El
ement 'p' at 0x7fe1b92b9868>, <Element 'p' at 0x7fe1b92b98b8>, <Element 'p' at 0
x7fe1b92b9908>, <Element 'p' at 0x7fe1b92b9958>, <Element 'p' at 0x7fe1b92b99a8>
, <Element 'p' at 0x7fe1b92b99f8>, <Element 'p' at 0x7fe1b92b9a48>, <Element 'p'
at 0x7fe1b92b9a98>]
allan@allan-ThinkPad-W510 ~/Documents/book/ch4/ch4/code $

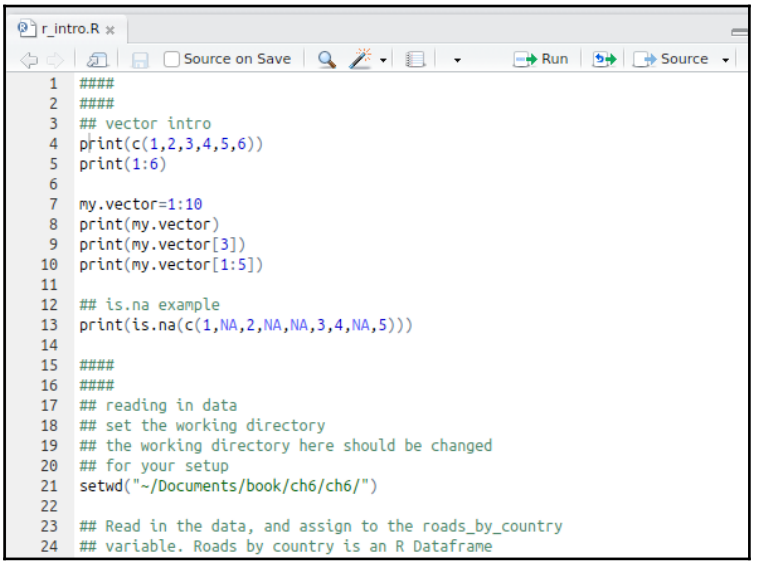
Chapter 05: Manipulating Text Data - An Introduction to Regular Expressions

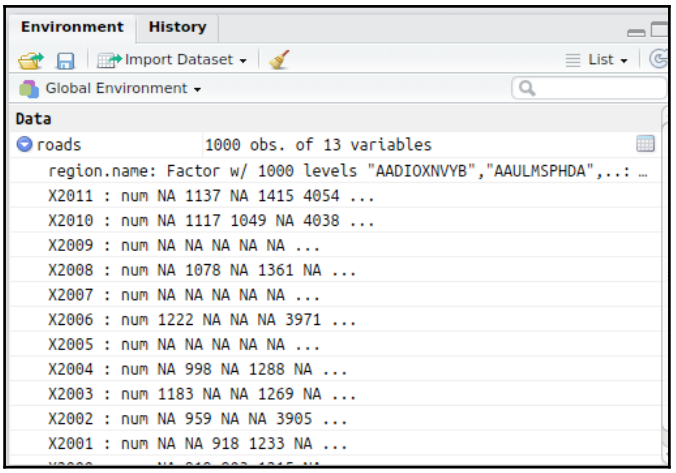


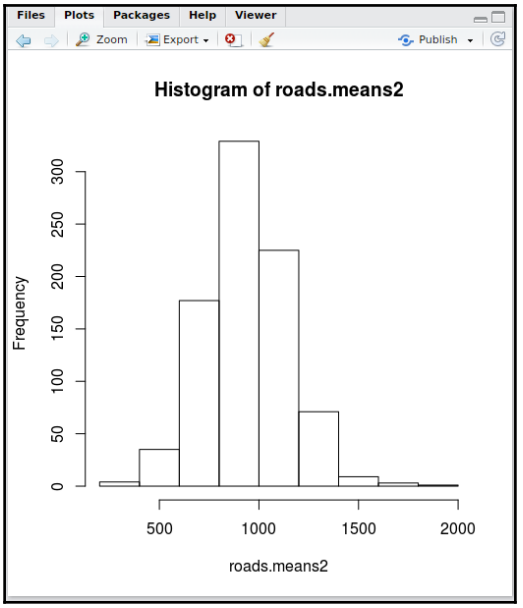

Terminal
52 howe street
20 center street
378 townsend avenue
2720 5th st
2630 5th street
2636 5th street
3032 jenn ave
79 bellerose drive
390 morris ave
41 baylor road
1771 25th st
2490 quebec avenue
404 winchester drive
10 wintergreen avenue
4100 knoll dr
550 cretin avenue
344 davis st
318 davis st
29 manning street
79 astor pl
637 3rd ave
23 gouverneur st
percent match: 0.4019730660820545
allan@allan-ThinkPad-W510 ~/Documents/book/ch5/ch5 $

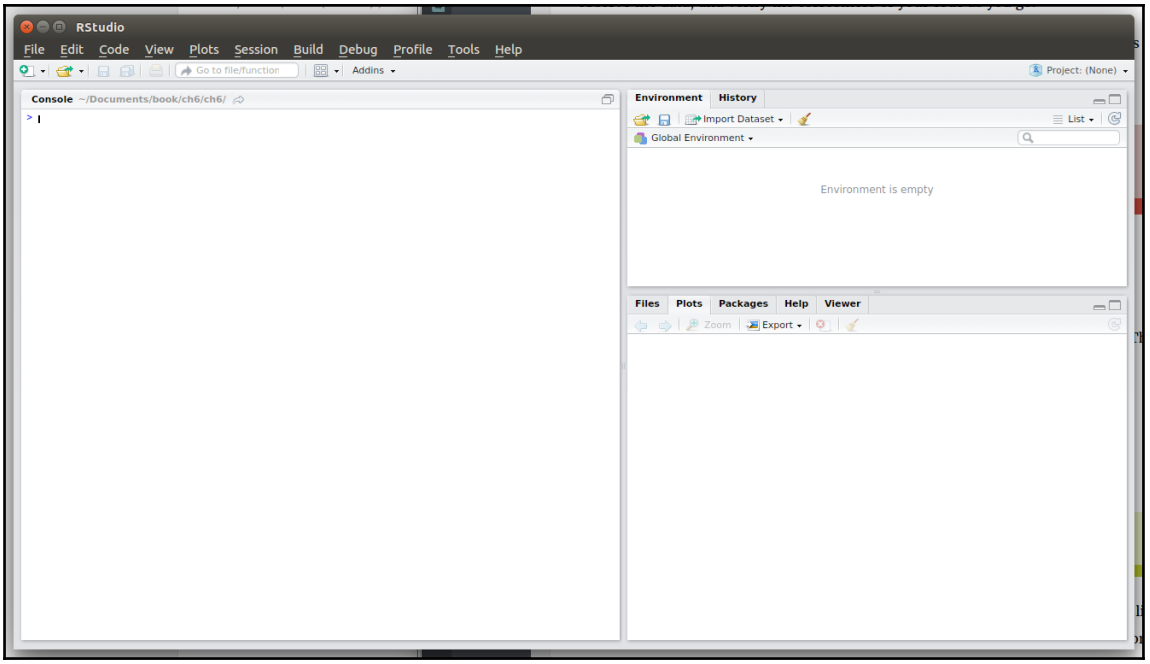

scf_street_name_data.csv - LibreOffice Calc
Liberation Sans
10
A1
created_at
C
D
E
summary
address
street_name
Street Signs
118 Church St Decatur, GA 30030, USA
church st
Code Enforcement (Illegal Dumping not in a street/Property Maintenance/Etc)
1104 Fullerton Ave Corona, CA 92879, USA
fullerton ave
Light
1653 Shatz St Las Vegas, Nevada
shatz st
Code Enforcement Issue
2456 Maple Street Atlanta, Georgia
maple street
SNOW RELATED
10 Field St New Haven, Connecticut
field st
Parking Enforcement
542 Chetwood St Oakland, CA 94610, USA
chetwood st
Signs / Bus Shelters / Pavement Markings
167 Hazel Street New Haven, Connecticut
hazel street
Yard and/or Structure Concern (High Grass, Junk, Broken Down Vehicles, Buildings, Sheds, Garages)
2552 Foraker Avenue Toledo, Ohio
foraker avenue
Litter/Shopping Carts
86 Copeland St Quincy, Massachusetts
copeland st
Asbestos Concern
2552 Foraker Avenue Toledo, Ohio
foraker avenue
Parking Enforcement
2824 Ritchie Street Oakland, California
ritchie street
Code Enforcement (Illegal Dumping not in a street/Property Maintenance/Etc)
315 Selkirk Dr Corona, CA 92881, USA
selkirk dr
Sidewalk
500 12th Street Oakland, California
12th street
Code Enforcement Police
252 Sully Rd Brockton, Massachusetts
sully rd
Customer drop damage
1010 Eastland Dr Lexington, KY 40505, USA
eastland dr
Pothole
2097 Como Ave St Paul, MN 55108, USA
como ave
Pothole
608 56th St Oakland, California
56th st
Traffic & Parking - Signs, Lights and Paint/Markings
112 Commercial Street Malden, MA
commercial street
Sidewalk Concern
9508 Frederick Road Ellicott City, Maryland
frederick road
Pot Hole
7839 12th Ave S Richfield/County, Minnesota
12th ave
scf_street_name_data
Find
Search Formatted Display String
Match Case
Sheet 1 of 1
Default
Sum=0
100%

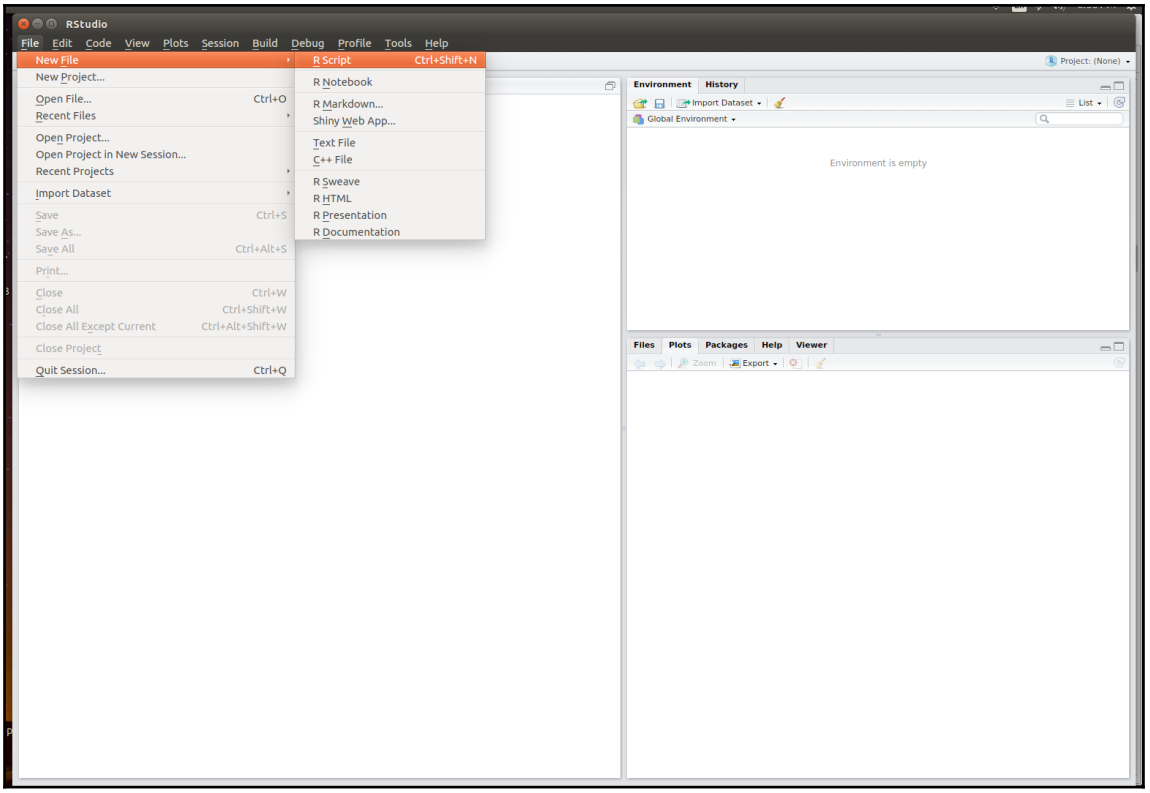
Chapter 06: Cleaning Numerical Data - An Introduction to R and RStudio

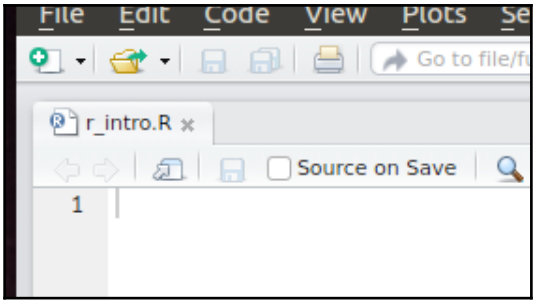


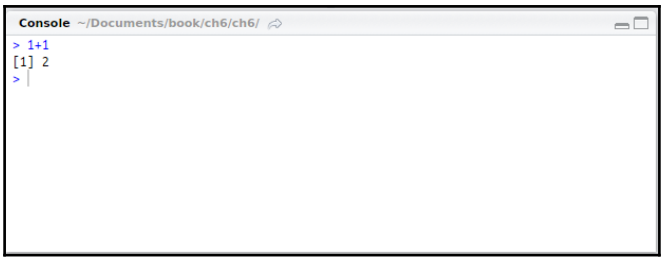

Environment
History
Import Dataset
List
Global Environment
Data
roads
1000 obs. of 13 variables
region.name: Factor w/ 1000 levels "AADIOXNVYB","AAULMSPHDA",..: …
X2011 : num NA 1137 NA 1415 4054 ...
X2010 : num NA 1117 1049 NA 4038 ...
X2009 : num NA NA NA NA NA ...
X2008 : num NA 1078 NA 1361 NA ...
X2007 : num NA NA NA NA NA ...
X2006 : num 1222 NA NA NA 3971 ...
X2005 : num NA NA NA NA NA ...
X2004 : num NA 998 NA 1288 NA ...
X2003 : num 1183 NA NA 1269 NA ...
X2002 : num NA 959 NA NA 3905 ...
X2001 : num NA NA 918 1233 NA ...

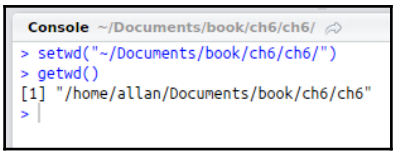

Files
Plots
Packages
Help
Viewer
Zoom
Export
Publish
Histogram of roads.means2
Frequency
0
50
100
150
200
250
300
500
1000
1500
2000
roads.means2

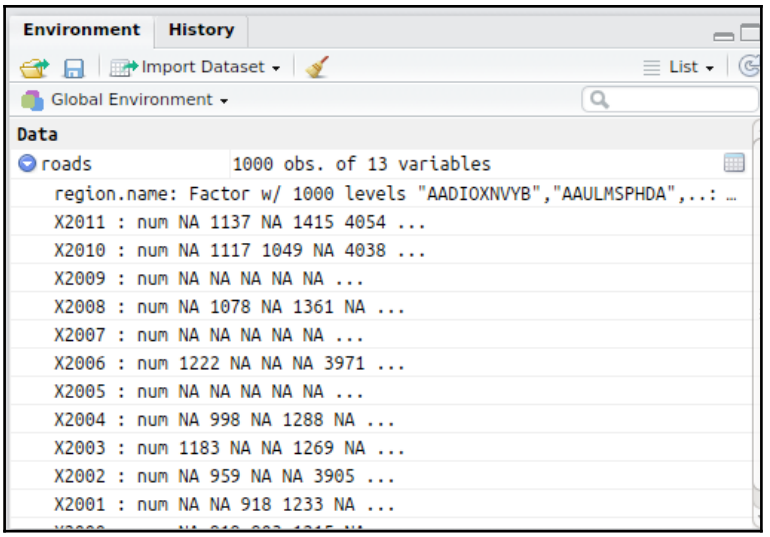

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Console ~/Documents/book/ch6/ch6/
Environment History
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Zoom
Export

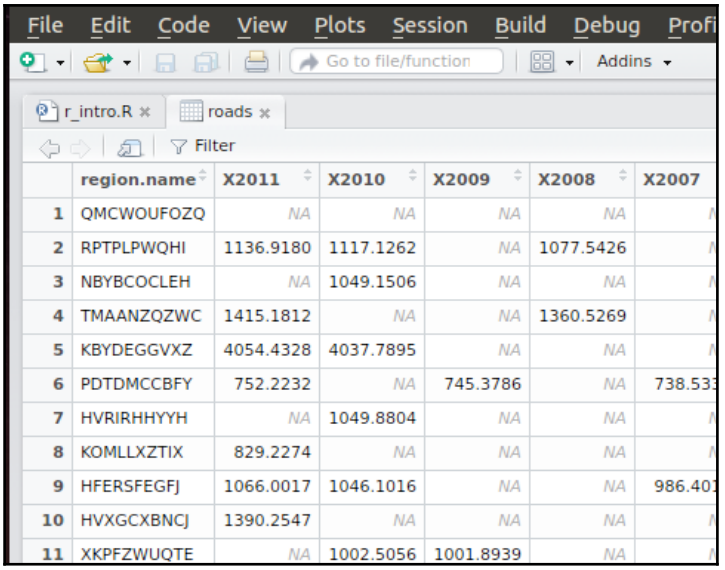

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
New File
New Project...
Open File... Ctrl+O
Recent Files
Open Project...
Open Project in New Session...
Recent Projects
Import Dataset
Save Ctrl+S
Save As...
Save All Ctrl+Alt+S
Print...
Close Ctrl+W
Close All Ctrl+Shift+W
Close All Except Current Ctrl+Alt+Shift+W
Close Project
Quit Session... Ctrl+Q
R Script Ctrl+Shift+N
R Notebook
R Markdown...
Shiny Web App...
Text File
C++ File
R Sweave
R HTML
R Presentation
R Documentation
Project: (None)
Environment History
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Zoom
Export


File Edit Code View Plots
Go to file/f
r_intro.R
Source on Save
1

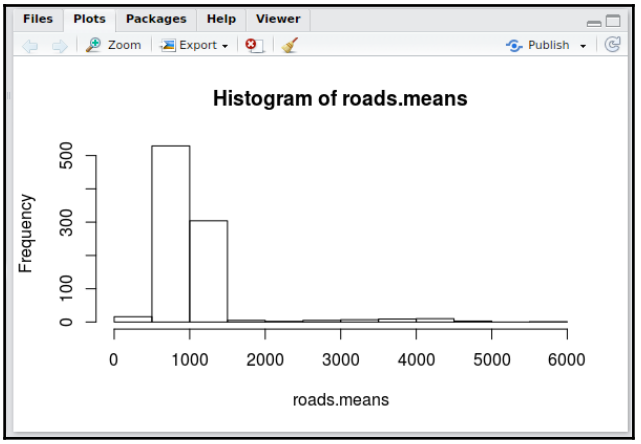

Console ~/Documents/book/ch6/ch6/
> 1+1
[1] 2
>

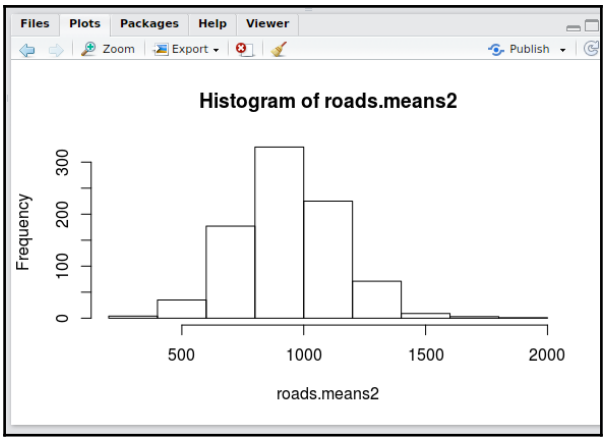

Console ~/Documents/book/ch6/ch6/
> setwd("~/Documents/book/ch6/ch6/")
> getwd()
[1] "/home/allan/Documents/book/ch6/ch6"
>


Environment
History
Import Dataset
List
Global Environment
Data
roads 1000 obs. of 13 variables
region.name: Factor w/ 1000 levels "AADIOXNVYB","AAULMSPHDA",..: …
X2011 : num NA 1137 NA 1415 4054 ...
X2010 : num NA 1117 1049 NA 4038 ...
X2009 : num NA NA NA NA NA ...
X2008 : num NA 1078 NA 1361 NA ...
X2007 : num NA NA NA NA NA ...
X2006 : num 1222 NA NA NA 3971 ...
X2005 : num NA NA NA NA NA ...
X2004 : num NA 998 NA 1288 NA ...
X2003 : num 1183 NA NA 1269 NA ...
X2002 : num NA 959 NA NA 3905 ...
X2001 : num NA NA 918 1233 NA ...



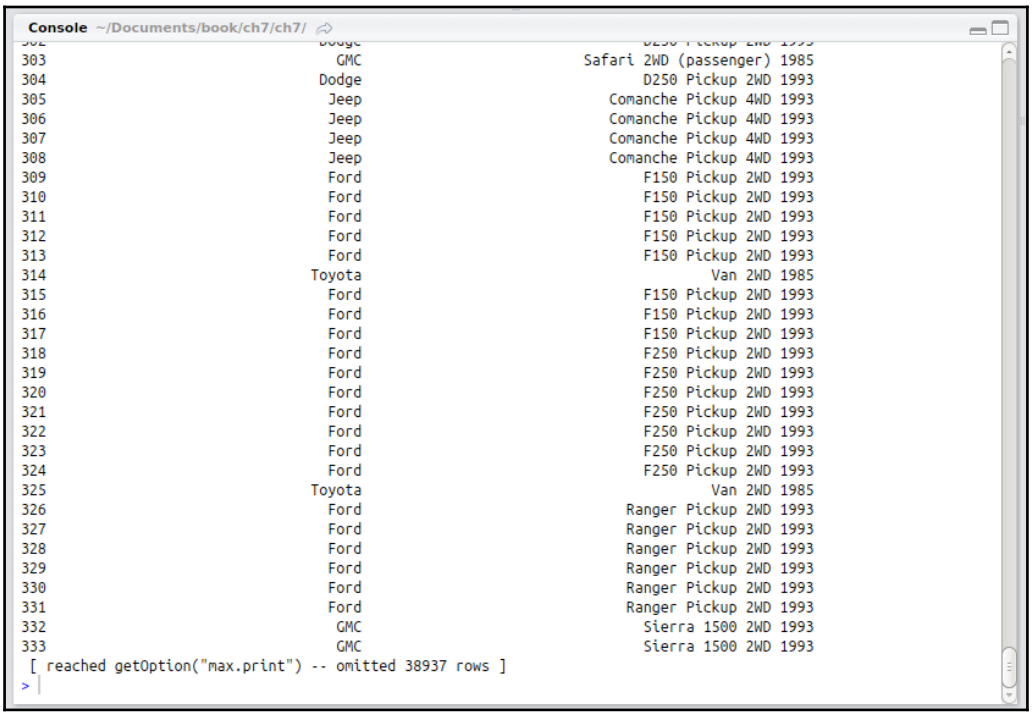
	region.name	X2011	X2010	X2009	X2008	X2007
1	QMCWOUFOZQ	NA	NA	NA	NA	
2	RPTPLPWQHI	1136.9180	1117.1262	NA	1077.5426	
3	NBYBCOCLEH	NA	1049.1506	NA	NA	
4	TMAANZQZWC	1415.1812	NA	NA	1360.5269	
5	KBYDEGGVXZ	4054.4328	4037.7895	NA	NA	
6	PDTDMCCBFY	752.2232	NA	745.3786	NA	738.53
7	HVRIRHHYYH	NA	1049.8804	NA	NA	
8	KOMLLXZTIX	829.2274	NA	NA	NA	
9	HFERSFEGFJ	1066.0017	1046.1016	NA	NA	986.40
10	HVXGCXBNCJ	1390.2547	NA	NA	NA	
11	XKPFZWUQTE	NA	1002.5056	1001.8939	NA	






```
+   return(as.numeric(x))
+ }
> roads.impute.na <- data.frame(
+   apply(roads.num3,2,impute,imputations=roads.means3)
+ )
> colSums(roads.impute.na)
   X2011    X2010    X2009    X2008    X2007    X2006    X2005    X2004    X2003    X2002    X2001    X2000
808648.3 807977.9 805344.1 802956.2 801247.2 799160.9 797380.7 795128.1 793213.5 790691.5 788098.0 786618.5
>
```

Chapter 07: Simplifying Data Manipulation with dplyr



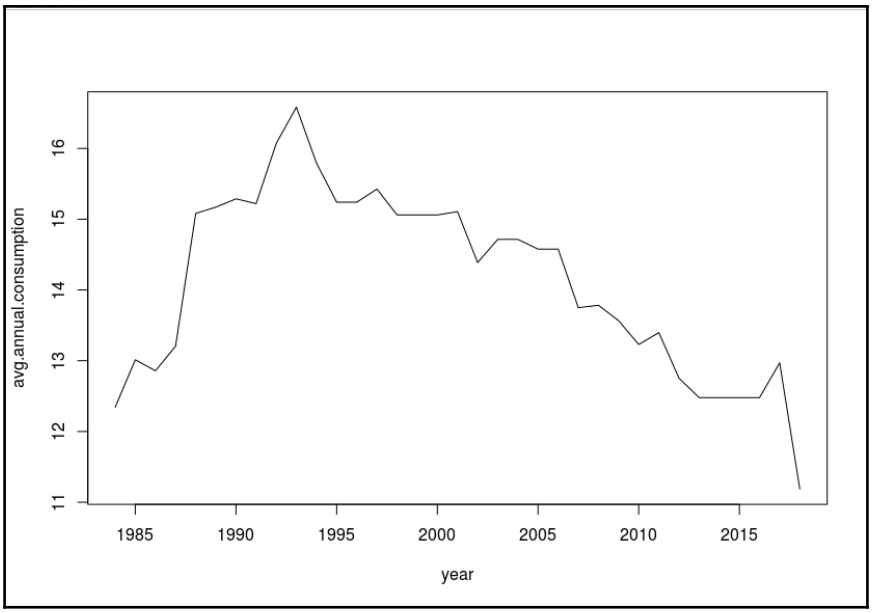
```
Console ~/Documents/book/ch7/ch7/
303                                GMC                        Safari 2WD (passenger) 1985
304                              Dodge                              D250 Pickup 2WD 1993
305                               Jeep                          Comanche Pickup 4WD 1993
306                               Jeep                          Comanche Pickup 4WD 1993
307                               Jeep                          Comanche Pickup 4WD 1993
308                               Jeep                          Comanche Pickup 4WD 1993
309                               Ford                              F150 Pickup 2WD 1993
310                               Ford                              F150 Pickup 2WD 1993
311                               Ford                              F150 Pickup 2WD 1993
312                               Ford                              F150 Pickup 2WD 1993
313                               Ford                              F150 Pickup 2WD 1993
314                             Toyota                                      Van 2WD 1985
315                               Ford                              F150 Pickup 2WD 1993
316                               Ford                              F150 Pickup 2WD 1993
317                               Ford                              F150 Pickup 2WD 1993
318                               Ford                              F250 Pickup 2WD 1993
319                               Ford                              F250 Pickup 2WD 1993
320                               Ford                              F250 Pickup 2WD 1993
321                               Ford                              F250 Pickup 2WD 1993
322                               Ford                              F250 Pickup 2WD 1993
323                               Ford                              F250 Pickup 2WD 1993
324                               Ford                              F250 Pickup 2WD 1993
325                             Toyota                                      Van 2WD 1985
326                               Ford                            Ranger Pickup 2WD 1993
327                               Ford                            Ranger Pickup 2WD 1993
328                               Ford                            Ranger Pickup 2WD 1993
329                               Ford                            Ranger Pickup 2WD 1993
330                               Ford                            Ranger Pickup 2WD 1993
331                               Ford                            Ranger Pickup 2WD 1993
332                                GMC                              Sierra 1500 2WD 1993
333                                GMC                              Sierra 1500 2WD 1993
 [ reached getOption("max.print") -- omitted 38937 rows ]
>
```

```
> print(vehicles.product.as.tibble)
# A tibble: 39,270 x 3
         make                model  year
       <fctr>              <fctr> <int>
 1 Alfa Romeo  Spider Veloce 2000  1985
 2    Ferrari          Testarossa  1985
 3      Dodge             Charger  1985
 4      Dodge B150/B250 Wagon 2WD  1985
 5     Subaru    Legacy AWD Turbo  1993
 6     Subaru              Loyale  1993
 7     Subaru              Loyale  1993
 8     Toyota             Corolla  1993
 9     Toyota             Corolla  1993
10     Toyota             Corolla  1993
# ... with 39,260 more rows
>
```

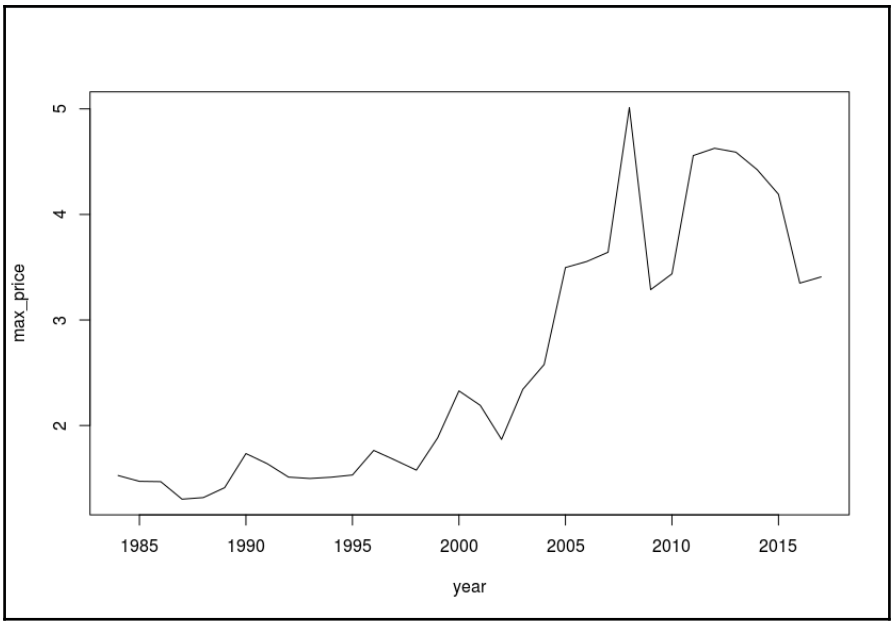
```
> print(vehicles.product.arranged)
# A tibble: 39,270 x 3
     make       model  year
   <fctr>      <fctr> <int>
 1  Acura 2.2CL/3.0CL  1997
 2  Acura 2.2CL/3.0CL  1997
 3  Acura 2.2CL/3.0CL  1997
 4  Acura 2.3CL/3.0CL  1998
 5  Acura 2.3CL/3.0CL  1998
 6  Acura 2.3CL/3.0CL  1998
 7  Acura 2.3CL/3.0CL  1999
 8  Acura 2.3CL/3.0CL  1999
 9  Acura 2.3CL/3.0CL  1999
10  Acura       2.5TL  1995
# ... with 39,260 more rows
```

```
> print(vehicles.camry)
# A tibble: 125 x 3
     make  model  year
   <fctr> <fctr> <int>
 1 Toyota  Camry  1993
 2 Toyota  Camry  1993
 3 Toyota  Camry  1993
 4 Toyota  Camry  1993
 5 Toyota  Camry  1994
 6 Toyota  Camry  1994
 7 Toyota  Camry  1994
 8 Toyota  Camry  1995
 9 Toyota  Camry  1995
10 Toyota  Camry  1995
# ... with 115 more rows
>
```

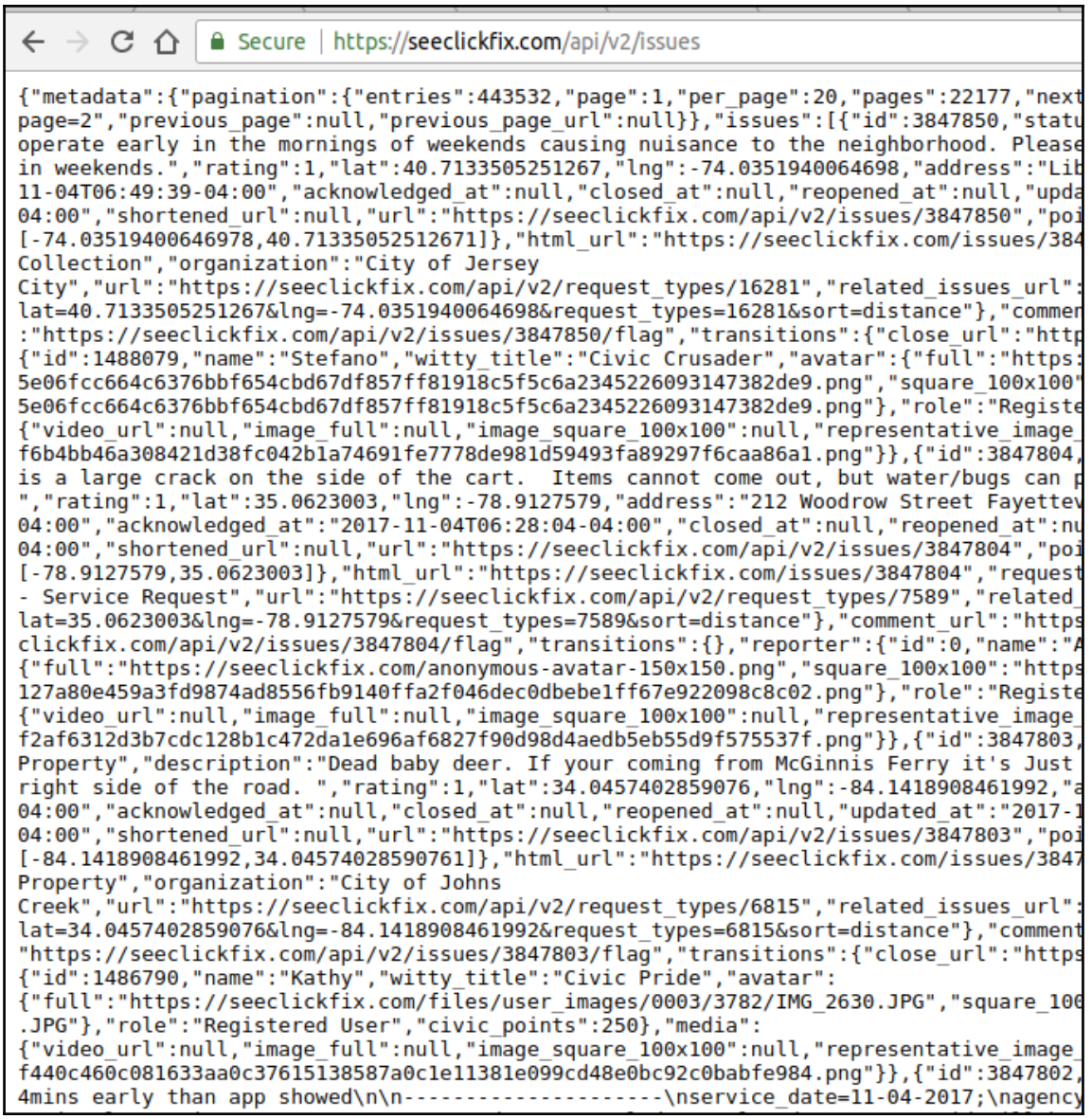
```
> print(camry.fuel_economy.by_year)
# A tibble: 35 x 2
    year avg.annual.consumption
   <int>                  <dbl>
 1  1984               12.34269
 2  1985               13.01336
 3  1986               12.85699
 4  1987               13.20553
 5  1988               15.08234
 6  1989               15.16997
 7  1990               15.28812
 8  1991               15.22018
 9  1992               16.07121
10  1993               16.58439
# ... with 25 more rows
>
```



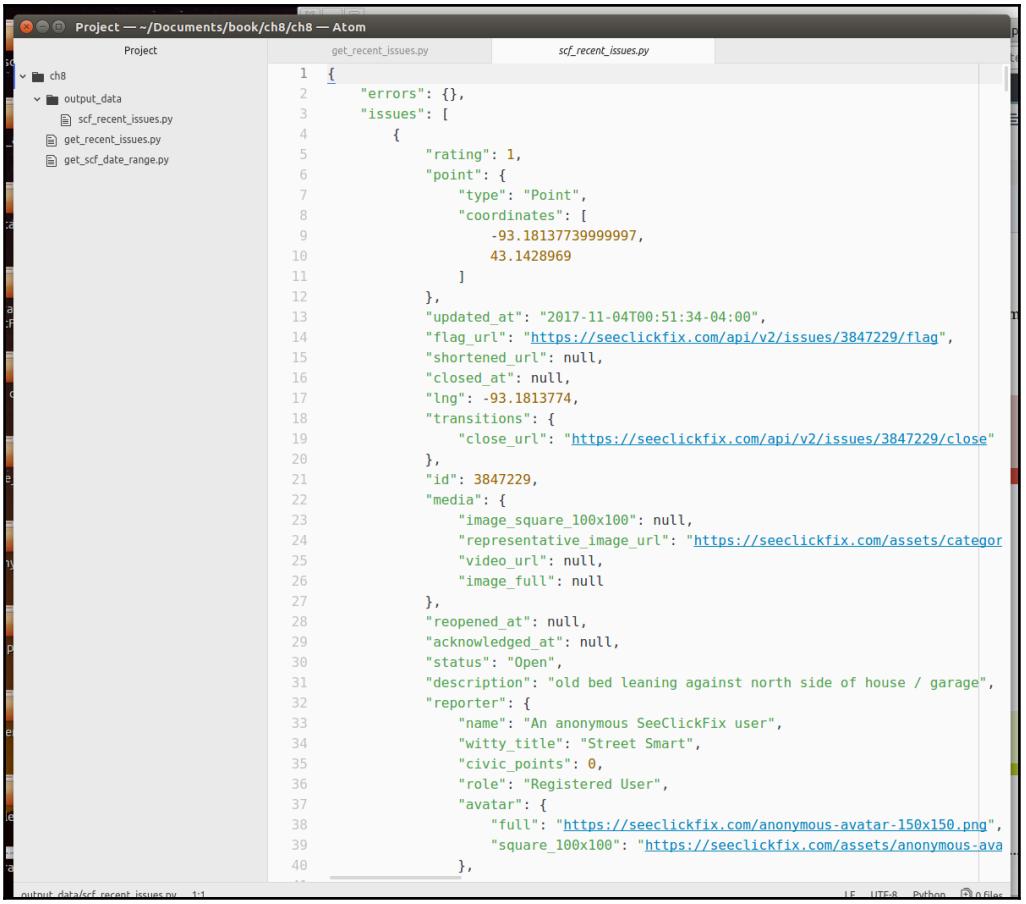

```
> print(gas_prices.by_year)
# A tibble: 34 x 2
    year max_price
   <int>     <dbl>
 1  1984     1.526
 2  1985     1.471
 3  1986     1.468
 4  1987     1.302
 5  1988     1.317
 6  1989     1.411
 7  1990     1.734
 8  1991     1.638
 9  1992     1.511
10  1993     1.498
# ... with 24 more rows
>
```


max_price
2
3
4
5
1985
1990
1995
2000
2005
2010
2015
year

Chapter 08: Getting Data from the Web

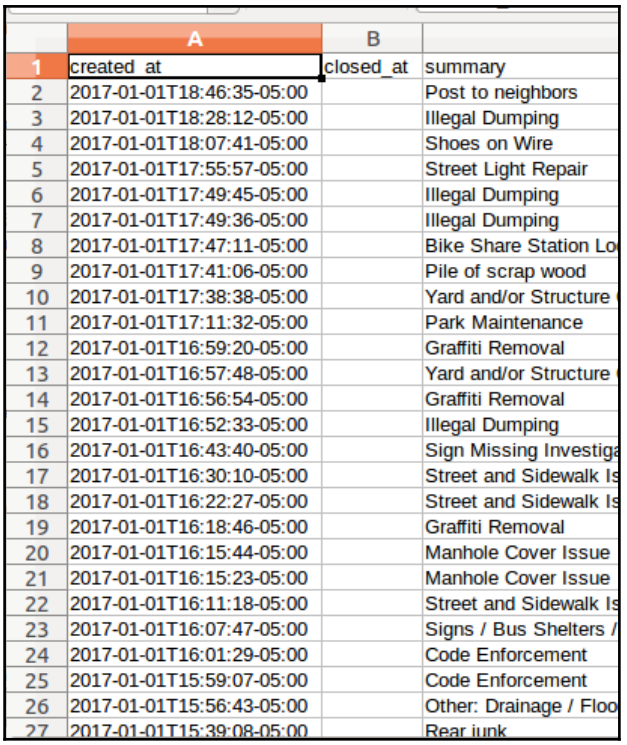



```
{"metadata":{"pagination":{"entries":443532,"page":1,"per_page":20,"pages":22177,"next
page=2","previous_page":null,"previous_page_url":null}},"issues":[{"id":3847850,"statu
operate early in the mornings of weekends causing nuisance to the neighborhood. Please
in weekends.","rating":1,"lat":40.7133505251267,"lng":-74.0351940064698,"address":"Lib
11-04T06:49:39-04:00","acknowledged_at":null,"closed_at":null,"reopened_at":null,"upda
04:00","shortened_url":null,"url":"https://seeclickfix.com/api/v2/issues/3847850","poi
[-74.03519400646978,40.71335052512671]},"html_url":"https://seeclickfix.com/issues/384
Collection","organization":"City of Jersey
City","url":"https://seeclickfix.com/api/v2/request_types/16281","related_issues_url":
lat=40.7133505251267&lng=-74.0351940064698&request_types=16281&sort=distance"},"commen
:"https://seeclickfix.com/api/v2/issues/3847850/flag","transitions":{"close_url":"http
{"id":1488079,"name":"Stefano","witty_title":"Civic Crusader","avatar":{"full":"https:
5e06fcc664c6376bbf654cbd67df857ff81918c5f5c6a2345226093147382de9.png","square_100x100"
5e06fcc664c6376bbf654cbd67df857ff81918c5f5c6a2345226093147382de9.png"},"role":"Registe
{"video_url":null,"image_full":null,"image_square_100x100":null,"representative_image_
f6b4bb46a308421d38fc042b1a74691fe7778de981d59493fa89297f6caa86a1.png"}},{"id":3847804,
is a large crack on the side of the cart.  Items cannot come out, but water/bugs can p
","rating":1,"lat":35.0623003,"lng":-78.9127579,"address":"212 Woodrow Street Fayettev
04:00","acknowledged_at":"2017-11-04T06:28:04-04:00","closed_at":null,"reopened_at":nu
04:00","shortened_url":null,"url":"https://seeclickfix.com/api/v2/issues/3847804","poi
[-78.9127579,35.0623003]},"html_url":"https://seeclickfix.com/issues/3847804","request
- Service Request","url":"https://seeclickfix.com/api/v2/request_types/7589","related_
lat=35.0623003&lng=-78.9127579&request_types=7589&sort=distance"},"comment_url":"https
clickfix.com/api/v2/issues/3847804/flag","transitions":{},"reporter":{"id":0,"name":"A
{"full":"https://seeclickfix.com/anonymous-avatar-150x150.png","square_100x100":"https
127a80e459a3fd9874ad8556fb9140ffa2f046dec0dbebe1ff67e922098c8c02.png"},"role":"Registe
{"video_url":null,"image_full":null,"image_square_100x100":null,"representative_image_
f2af6312d3b7cdc128b1c472da1e696af6827f90d98d4aedb5eb55d9f575537f.png"}},{"id":3847803,
Property","description":"Dead baby deer. If your coming from McGinnis Ferry it's Just
right side of the road. ","rating":1,"lat":34.0457402859076,"lng":-84.1418908461992,"a
04:00","acknowledged_at":null,"closed_at":null,"reopened_at":null,"updated_at":"2017-1
04:00","shortened_url":null,"url":"https://seeclickfix.com/api/v2/issues/3847803","poi
[-84.1418908461992,34.04574028590761]},"html_url":"https://seeclickfix.com/issues/3847
Property","organization":"City of Johns
Creek","url":"https://seeclickfix.com/api/v2/request_types/6815","related_issues_url":
lat=34.0457402859076&lng=-84.1418908461992&request_types=6815&sort=distance"},"comment
"https://seeclickfix.com/api/v2/issues/3847803/flag","transitions":{"close_url":"https
{"id":1486790,"name":"Kathy","witty_title":"Civic Pride","avatar":
{"full":"https://seeclickfix.com/files/user_images/0003/3782/IMG_2630.JPG","square_100
.JPG"},"role":"Registered User","civic_points":250},"media":
{"video_url":null,"image_full":null,"image_square_100x100":null,"representative_image_
f440c460c081633aa0c37615138587a0c1e11381e099cd48e0bc92c0babfe984.png"}},{"id":3847802,
4mins early than app showed\n\n----------\nservice_date=11-04-2017;\nagency
```



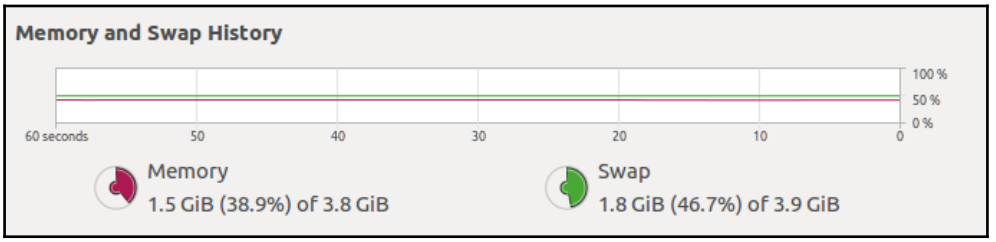
- **after=**`:time` - must be a timestamp in ISO 8601 format: YYYY-MM-DDTHH:MM:SSZ
- **before=**`:time` - must be a timestamp in ISO 8601 format: YYYY-MM-DDTHH:MM:SSZ

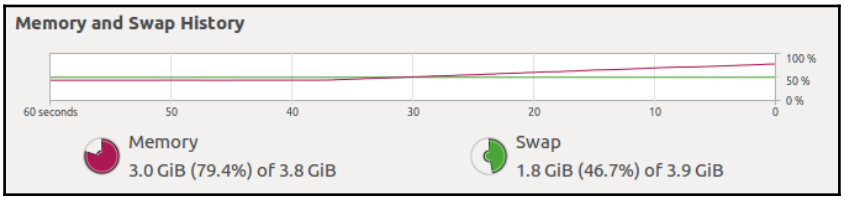
- **page=**`:page_number` - number of the page to return, default: 1

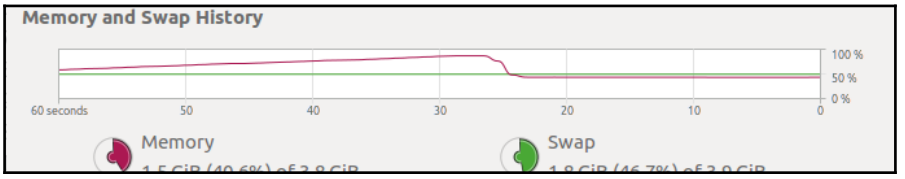


	A	B	
1	created_at	closed_at	summary
2	2017-01-01T18:46:35-05:00		Post to neighbors
3	2017-01-01T18:28:12-05:00		Illegal Dumping
4	2017-01-01T18:07:41-05:00		Shoes on Wire
5	2017-01-01T17:55:57-05:00		Street Light Repair
6	2017-01-01T17:49:45-05:00		Illegal Dumping
7	2017-01-01T17:49:36-05:00		Illegal Dumping
8	2017-01-01T17:47:11-05:00		Bike Share Station Lo
9	2017-01-01T17:41:06-05:00		Pile of scrap wood
10	2017-01-01T17:38:38-05:00		Yard and/or Structure
11	2017-01-01T17:11:32-05:00		Park Maintenance
12	2017-01-01T16:59:20-05:00		Graffiti Removal
13	2017-01-01T16:57:48-05:00		Yard and/or Structure
14	2017-01-01T16:56:54-05:00		Graffiti Removal
15	2017-01-01T16:52:33-05:00		Illegal Dumping
16	2017-01-01T16:43:40-05:00		Sign Missing Investiga
17	2017-01-01T16:30:10-05:00		Street and Sidewalk Is
18	2017-01-01T16:22:27-05:00		Street and Sidewalk Is
19	2017-01-01T16:18:46-05:00		Graffiti Removal
20	2017-01-01T16:15:44-05:00		Manhole Cover Issue
21	2017-01-01T16:15:23-05:00		Manhole Cover Issue
22	2017-01-01T16:11:18-05:00		Street and Sidewalk Is
23	2017-01-01T16:07:47-05:00		Signs / Bus Shelters /
24	2017-01-01T16:01:29-05:00		Code Enforcement
25	2017-01-01T15:59:07-05:00		Code Enforcement
26	2017-01-01T15:56:43-05:00		Other: Drainage / Floo
27	2017-01-01T15:39:08-05:00		Rear junk

Chapter 09: Working with Large Datasets








```
> use weather
switched to db weather
> db.records.count()
94608000
> db.records.findOne()
{
        "_id" : ObjectId("5a001e38b4b746a7889b644a"),
        "date" : "1980-01-01 00:00:00",
        "temperature" : 2.429978875842491,
        "is_cloudy" : 1,
        "is_sunny" : 0
}
>
```

```
2491, "is_cloudy" : 1, "is_sunny" : 0 }

136, "is_cloudy" : 0, "is_sunny" : 0 }
4306, "is_cloudy" : 1, "is_sunny" : true }
3997, "is_cloudy" : 1, "is_sunny" : true }
3853, "is_cloudy" : 1, "is_sunny" : true }
91754, "is_cloudy" : 0, "is_sunny" : 0 }
51694, "is_cloudy" : 0, "is_sunny" : true }
86484, "is_cloudy" : 0, "is_sunny" : true }
8906, "is_cloudy" : 1, "is_sunny" : true }
45535, "is_cloudy" : 1, "is_sunny" : 0 }
64791, "is_cloudy" : 1, "is_sunny" : 0 }
49577, "is_cloudy" : 0, "is_sunny" : true }
5621, "is_cloudy" : 0, "is_sunny" : true }
24433, "is_cloudy" : 1, "is_sunny" : true }
95258, "is_cloudy" : 0, "is_sunny" : true }
7857, "is_cloudy" : 1, "is_sunny" : 0 }
43853, "is_cloudy" : 1, "is_sunny" : true }
56436, "is_cloudy" : 0, "is_sunny" : true }
68617, "is_cloudy" : 0, "is_sunny" : 0 }
```

Index