Teaching data science with puzzles

useR! 2019
Irene Steves
i_steves isteves

bit.ly/ds-puzzles
--- Day 11: Hex Ed ---

Crossing the bridge, you've barely reached the other side of the stream when a program comes up to you, clearly in distress. "It's my child process," she says, "he's gotten lost in an infinite grid!"

Fortunately for her, you have plenty of experience with infinite grids.

Unfortunately for you, it's a hex grid.

The hexagons ("hexes") in this grid are aligned such that adjacent hexes can be found to the north, northeast, southeast, south, southwest, and northwest:

You have the path the child process took. Starting where he started, you need to determine the fewest number of steps required to reach him. (A "step" means to move from the hex you are in to any adjacent hex.)

For example:

- ne,ne,ne is 3 steps away.
- ne,ne,sw,sw is 0 steps away (back where you started).
- ne,ne,s,s is 2 steps away (se,se).
- se,sw,se,sw,se is 3 steps away (s,s,sw).

To begin, get your puzzle input.

Answer: [Submit]
Our sponsors help make Advent of Code possible:

Novetta - Unleash your imagination. Innovate at Novetta.

I solved these with R, but boy was it clunky!
Let’s make puzzles that highlight what R/the tidyverse are good at!
Bite-sized puzzles that focus on core data science skills as championed by the tidyverse set of packages

[still unreleased!]
SOOTHSAYER. Beware the ides of March.
CAESAR. What man is that?
BRUTUS. A soothsayer bids you beware the ides of March.
CAESAR. Set him before me; let me see his face.
CASSIUS. Fellow, come from the throng; look upon Caesar.
CAESAR. What say'st thou to me now? Speak once again.
SOOTHSAYER. Beware the ides of March.
CAESAR. He is a dreamer; let us leave him.
Wrangling
Web-based experience

Tidies of March

Select puzzle:

11_sandwiches

Select user ID:

1

The little sandwich store around the corner makes the best sandwiches! It's an adventure every time you go there—you can get everything from classics like Italian beef sandwiches to more exciting choices like Fufu, Muffler, and Kokorec sandwiches.

Unfortunately, they're spending so much on ingredients that they can't turn a profit. They've decided to cut their selection and only focus on their best-selling sandwich.

They've collected data on the favorite sandwiches among customers that came into the store in the last month. Most people ended up listing several sandwiches as their favorites (in no particular order), so the data looks like this:

<table>
<thead>
<tr>
<th>names</th>
<th>sandwiches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>Denver; Torta; Torta ahogada; Barbecue</td>
</tr>
<tr>
<td>Abigail</td>
<td>BLT; Pita; Primanti; Ice cream; Choripan</td>
</tr>
<tr>
<td>Adam</td>
<td>Corned beef; Montealito; Cheesesteak; Tripeata; Dagwood; Jambon-beurre</td>
</tr>
<tr>
<td>Alexa</td>
<td>Mortadella; Dagwood</td>
</tr>
<tr>
<td>Alexandria</td>
<td>Slider; Beuncht met mailejes; Chicken salad</td>
</tr>
<tr>
<td>Ana</td>
<td>Fried brain; Polish boy; Vegetables; Pudgy Pie; Dagwood</td>
</tr>
</tbody>
</table>

In this sample, the Dagwood sandwich is the most popular.

In the data provided, what is the most popular sandwich among the sandwich customers?

Click download to get your unique puzzle input (csv). If you need extra help, click hint for a list of useful functions that are relevant to this puzzle.

Download Hint

Your solution:

Submit

Language & platform agnostic
R-mediated experience → Workflow
initialize_puzzles(".")
New R Project
Tidies of March 2019

Table of Contents

1. pets
<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>01_data.xlsx</td>
<td>10.2 KB</td>
<td>Jan 16, 2019, 11:07 F</td>
</tr>
<tr>
<td>01_soln.R</td>
<td>297 B</td>
<td>Jan 16, 2019, 11:07 F</td>
</tr>
<tr>
<td>01_text.Rmd</td>
<td>1.7 KB</td>
<td></td>
</tr>
</tbody>
</table>

three principles for (file) names

- machine readable
- human readable
- plays well with default ordering
Knittable .R file

```
#' ---
#' title: Pets
#' ---

#' Use _Ctrl (Cmd) + Shift + K_ to render this file
#
#+ r setup, include = FALSE
options(tidyverse.quiet = TRUE)

#+ r
library(tidyverse)
library(here)
data_path <- here::here('01_pets', '01_data.xlsx')

# YOUR SOLUTION CODE HERE
```

Paths that work in the console & when rendered

Omit tidyverse messages from html output
The neighborhood deli makes amazing sandwiches--from classics like BLTs to dessert sandwiches like Fluffernutters. Since many of their specialty ingredients keep going bad, they've decided to cut their selection and focus on their best-selling sandwich.
To help with the decision, the storeowners collected data on their customers’ favorites. Most people listed several varieties (in no particular order). Here’s a sample of the data:

<table>
<thead>
<tr>
<th>names</th>
<th>sandwiches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>Denver; BLT; Torta ahogada; Barbecue</td>
</tr>
<tr>
<td>Abigail</td>
<td>BLT; Ftira; Primanti; Ice cream; Choripán</td>
</tr>
<tr>
<td>Adam</td>
<td>Corned beef; Montadito; Cheesesteak; Tripleta; Dagwood; Jambon-beurre</td>
</tr>
<tr>
<td>Alexa</td>
<td>Dagwood; Mortadella</td>
</tr>
<tr>
<td>Alexandria</td>
<td>Slider; Beschuit met muisjes; Chicken salad</td>
</tr>
<tr>
<td>Ana</td>
<td>Fried brain; Polish boy; Vegetable; Pudgy Pie; Dagwood</td>
</tr>
</tbody>
</table>

In this sample, the Dagwood sandwich is the most popular.

In the full dataset, **what is the most popular sandwich among the customers?**
<table>
<thead>
<tr>
<th>names</th>
<th>sandwiches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>Denver; BLT; Torta ahogada; Barbecue</td>
</tr>
<tr>
<td>Abigail</td>
<td>BLT; Ftira; Primanti; Ice cream; Choripán</td>
</tr>
<tr>
<td>Adam</td>
<td>Corned beef; Montadito; Cheesesteak; Tripleta; Dagwood; Jambon-beurre</td>
</tr>
<tr>
<td>Alexa</td>
<td>Dagwood; Mortadella</td>
</tr>
<tr>
<td>Alexandria</td>
<td>Slider; Beschuit met muisjes; Chicken salad</td>
</tr>
<tr>
<td>Ana</td>
<td>Fried brain; Polish boy; Vegetable; Pudgy Pie; Dagwood</td>
</tr>
</tbody>
</table>

In this sample, the **Dagwood** sandwich is the most popular.
## A tibble: 6 x 2
## names    sandwiches
## <chr>    <chr>
## 1 Abby    Denver; BLT; Torta ahogada; Barbecue
## 2 Abigail BLT; Ftira; Primanti; Ice cream; Choripán
## 3 Adam    Corned beef; Montadito; Cheesesteak; Tripleta; Dagwood; Jamb
## 4 Alexa   Dagwood; Mortadella
## 5 Alexandria Slider; Beschuit met muisjes; Chicken salad
## 6 Ana     Fried brain; Polish boy; Vegetable; Pudgy Pie; Dagwood
```r
separate_rows(sandwiches, sep = "; ")
```

## A tibble: 25 x 2
##     names   sandwiches
##   <chr>     <chr>
## 1  Abby   Denver
## 2  Abby   BLT
## 3  Abby Torta ahogada
## 4  Abby  Barbecue
## 5 Abigail BLT
## 6 Abigail Ftira
## 7 Abigail Primanti
## 8 Abigail Ice cream
## 9 Abigail Choripán
##10  Adam Corned beef
## # ... with 15 more rows
## A tibble: 22 x 2
##
##  sandwiches n
##  <chr>     <int>
## 1 Dagwood  3
## 2 BLT      2
## 3 Barbecue 1
## 4 Beschuit met muisjes 1
## 5 Cheesesteak 1
## 6 Chicken salad 1
## 7 Choripán 1
## 8 Corned beef 1
## 9 Denver 1
##10 Fried brain 1
## # ... with 12 more rows
tidyverse
Beyond the tidyverse

Consistent and parseable names

Test cases

Self-contained code

Projects & version control

R

git

rmarkdown

repex
Thank you!
Irene Steves
i_stevesssteves
bit.ly/ds-puzzles