# RPB Engineering | Coding Challenge

Thanks for applying to Rad Power Bikes. We really enjoyed chatting with you, and we'd love to see more of what you can do. To better understand your style, we'd love to see how you would handle the construction of a sample data pipeline for our organization in your language of choosing.

## Story

Given `N` serial numbers found in .csv files (ex: `bikes_received_20207001_010355.csv` for bikes received on July 1, 2020 at 01:13:55 AM) that are placed at intervals by vendors in a shared directory, our program will extract the daily file,decode it, and place it into our data warehouse.

### Serial Codes:

For Example, if file `bikes_received_20207001_010355.csv` contains

```
VC719V1004706,SC420V3000007,7ABCM10000123
```

or…

```
VC719V1004706
SC420V3000007
7ABCM10000123
```

the data returned will be the following:

| model | model year | manufactured month | manufactured year | version | unique | dt_received |
|---|---|---|---|---|---|---|
| Runner | 2020 | July | 19 | 1 | 004706 | 2020-07-01T01:03:55.000000 |
| RadCity Stepthru | 2020 | April | 20 | 3 | 000007 | 2020-07-01T01:03:55.000000 |
| Unknown | Unknown | Unknown | Unknown | 10 | 000123 | 2020-07-01T01:03:55.000000 |

Fortunately, the team upstairs has a spread sheet (🤮 ) that explains the method to the format madness:

| Example of New serial number | Bike model | Model year code | Month Codes | Year Manufactured (last 2 digits) in YY format (2017 would be 17 and 2018 would be 18) | Assembly Plant Code | Version of the bike (Revisions in 1, 2, 3) | serial number 6 number |
|---|---|---|---|---|---|---|---|
| RB719F1000001 | R | B | 7 | 19 | F | 1 | 000001 |

| Example of New serial number | Bike model | Model year code | Month Codes | Year Manufactured (last 2 digits) in YY format (2017 would be 17 and 2018 would be 18) | Assembly Plant Code | Version of the bike (Revisions in 1, 2, 3) | serial numbe 6 numbe |
|---|---|---|---|---|---|---|---|
| HB719F1000001 | H | B | 7 | 19 | F | 1 | 000001 |
| SB919F1000001 | S | B | 9 | 19 | F | 1 | 000001 |

| | Model Codes | | Month Code | | Model Year Code | | Factory Code | |
|---|---|---|---|---|---|---|---|---|
| | R | RadRover | 1 | January | A | 2018 | F | FactoryF |
| | M | RadMini | 2 | February | B | 2019 | V | FactoryV |
| | W | RadWagon | 3 | March | C | 2020 | | |
| | 6 | RadCity 16 | 4 | April | D | 2021 | | |
| | 9 | RadCity 19 | 5 | May | | | | |
| | S | RadCity Stepthru | 6 | June | | | | |
| | B | RadBurro | 7 | July | | | | |
| | H | RadRhino | 8 | August | | | | |
| | C | Large Cargo Box | 9 | September | | | | |
| | K | Small Cargo Box | O | October | | | | |
| | P | Pedicab | N | November | | | | |
| | F | Flatbed | D | December | | | | |
| | T | Truckbed | | | | | | |
| | N | Insulated Cargo Box | | | | | | |
| | V | Runner | | | | | | |

# Requirements

There are a few requirements by which you will need to abide.

- You must provide a **backend** that:
  - On a daily cadence, checks a directory and parses only net new files

- Checks for and avoids data duplication
- Does not delete any data from the database
- Feel free to use whichever database system you would like as a data warehouse

## Icing on the 🍰

- Archive parsed files into a new directory, then clear the current directory so as to keep things clean for the following day