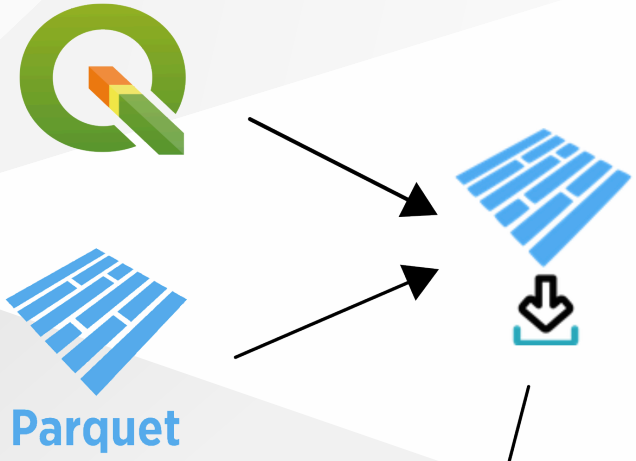
Geoparquet



GeoParquet Downloader

```r
# Load required packages

library(arrow)
library(sf)
library(sfarrow)
library(dplyr)

```

Parquet

10:	200:	1004:	237738.227361176946
22:	200:	1004:	257573.399901173348
28:	200:	2004:	393573.333311173874
16:	100:	1002:	135573.667341115474
19:	100:	4974:	449051.369971156664
45:	100:	1955:	463158.313331155264
47:	200:	2864:	445663.213581133659
16:	100:	2054:	991004.399201128667
17:	100:	2007:	393892.349671138368
25:	100:	1974:	995751.983971176023
24:	200:	4971:	395364.397701195954

Marp

Introduction

- Geoparquet is a columnar storage format for geographic data
- It is based on the Apache Parquet format
- It is designed to be efficient for both storage and retrieval of geographic data
- There are now more than 20 different tools and libraries in 6 different languages that support GeoParquet, including R (geoarrow, sfarrow), Python (Geopandas, Pyarrow, Fiona), and Julia (GeoParquet.jl).

Use Geoparquet with R

```
library(sfarrow)
library(sf)
```

First steps

- Read a geoparquet file using the `st_read_parquet` function from the `sfarrow` package
- Write a geoparquet file using the `st_write_parquet` function from the `sfarrow` package

Example

```
data <- st_read_parquet('data/example.parquet')
plot(sf::st_geometry(data))
st_write_parquet(data, 'data/example.parquet')
```

Implementation Details

- The `st_read_parquet` function reads a geoparquet file and returns a `sf` object.
- The `st_write_parquet` function writes a `sf` object to a geoparquet file.
- The `sfarrow` is a package for reading and writing Parquet and Feather files with `sf` objects using `arrow` in R.
- The `sf` package is used to work with spatial data in R.

Use dplyr with Geoparquet

```
library(dplyr)
data <- st_read_parquet('data/example.parquet')
data <- data %>%
  filter(class_type == 'Primary') %>%
  mutate(id = row_number()) %>%
  select(geometry, id)
st_write_parquet(data, 'data/filtered_example.parquet')
```

Other example:

```
groups <- data %>%
    filter(subtype == 'Type1') %>%
    group_by(class) %>%
    summarize(Total = n_distinct(row_number())) %>%
    st_drop_geometry() %>%
    write.csv('data/table_summary.csv')
```

Use sf with Geoparquet

```
library(sf)
data <- st_read_parquet('data/example.parquet')
data_buffer <- data %>%
  mutate(length = st_length(geometry)) %>%
  st_buffer(100)

plot(sf::st_geometry(data_buffer))

st_write_parquet(data, 'data/buffer_example.parquet')
```

Use tidyverse with Geoparquet

```
library(tidyverse)
data <- st_read_parquet('data/example.parquet')
data <- data %>%
  ggplot() +
  geom_sf() +
  ggtitle('Example') +
  theme(plot.title = element_text(hjust = 0.5)) +
  coord_sf() +
  theme_minimal()

ggsave('data/example.png', data, width = 10, height = 10, dpi = 300)
```