Introduction to NumPy

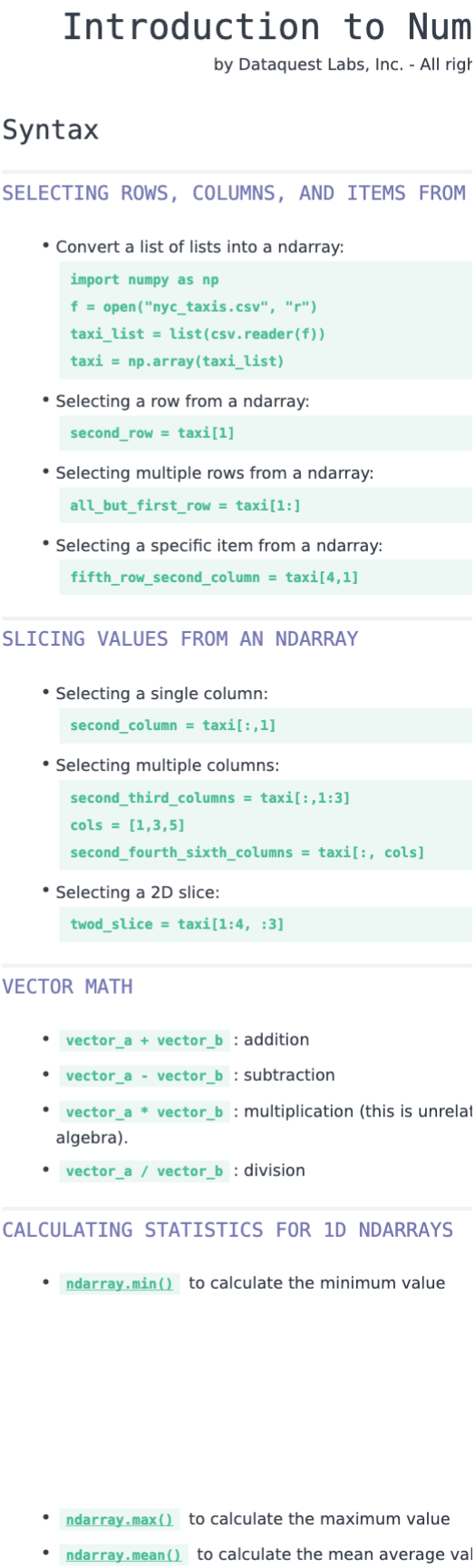
Next Steps

In this lesson, we learned the following:

- How vectorization speeds up our code
- About n-dimensional arrays, and NumPy's ndarrays
- How to select specific items, rows, columns, 1D slices, and 2D slices from ndarrays
- How to apply simple calculations to entire ndarrays
- How to use vectorized methods to perform calculations across either axis of ndarrays

In the next lesson, we'll continue to work with the NYC taxi data as we learn about **Boolean indexing**, one of the most powerful tools when working with data in NumPy and pandas.

Finish



Introduction to Num

by Dataquest Labs, Inc. - All rigl

Syntax

SELECTING ROWS, COLUMNS, AND ITEMS FROM

- Convert a list of lists into a ndarray:

```
import numpy as np
f = open("nyc_taxis.csv", "r")
taxi_list = list(csv.reader(f))
taxi = np.array(taxi_list)
```

- Selecting a row from a ndarray:

```
second_row = taxi[1]
```

- Selecting multiple rows from a ndarray:

```
all_but_first_row = taxi[1:]
```

- Selecting a specific item from a ndarray:

```
fifth_row_second_column = taxi[4,1]
```

SLICING VALUES FROM AN NDARRAY

- Selecting a single column:

```
second_column = taxi[:,1]
```

- Selecting multiple columns:

```
second_third_columns = taxi[:,1:3]
cols = [1,3,5]
second_fourth_sixth_columns = taxi[:, cols]
```

- Selecting a 2D slice:

```
twod_slice = taxi[1:4, :3]
```

VECTOR MATH

- `vector_a + vector_b` : addition
- `vector_a - vector_b` : subtraction
- `vector_a * vector_b` : multiplication (this is unrelat algebra).
- `vector_a / vector_b` : division

CALCULATING STATISTICS FOR 1D NDARRAYS

- `ndarray.min()` to calculate the minimum value
- `ndarray.max()` to calculate the maximum value
- `ndarray.mean()` to calculate the mean average val