A summarization of the data:

Taking the median values for each feature, we can see that this person’s median daily distance travelled is 3.68 miles per day and the person median climb per day is 7 floors. The total activity adds up to only 16.3 hours, so the person must only be wearing the Fitbit while awake. That leaves 7.68 hours for sleep if he or she takes it off right before going to sleep and puts it on right after waking up. The person’s median sedentary time is 12.5 hours per day. The majority of activity is the 3.38 hours of active-light activity.

So likely the person has a desk job, but works out, maybe also walking into and out of work and taking the stairs. He or she could work on the third floor of the building, though the floors climbed could be done at the gym or at home.

Details in data:

The raw data as we first read it contained 17 features and 247 rows of data.

Int64Index: 247 entries, 0 to 246

Data columns (total 17 columns):

index 247 non-null datetime64[ns]

activity\_calories 225 non-null object

distance 225 non-null object

floors 225 non-null object

mins\_active\_high 225 non-null object

mins\_active\_light 225 non-null object

mins\_active\_med 225 non-null object

mins\_sedentary 225 non-null object

steps 225 non-null object

total\_calories 225 non-null object

calories 247 non-null object

carbs 247 non-null object

fat 247 non-null object

fiber 247 non-null object

protien 247 non-null object

sodium 247 non-null object

water 247 non-null object

dtypes: datetime64[ns](1), object(16)

The last twenty-two days contained only zeros or NaN values. We dropped the rows containing NaN values and that dropped those rows. Additionally, ten features ('calories', 'carbs', 'fat', 'fiber', 'protein', 'sodium', and 'water') had primarily only the default zeroes for values, so we dropped those columns as well.

All of the data was numeric, but, of course, it read in as objects. We took the commas out of the numbers and converted them to float type. We then set the index to the date, which was the original index.

After those conversions, all 9 features were dense ftypes, meaning they contain mostly non-zero values.

After all that manipulation, we the number of zero values we had in each feature were:  
activity\_calories 9

distance 12

floors 31

mins\_active\_high 51

mins\_active\_light 9

mins\_active\_med 46

mins\_sedentary 0

steps 12

total\_calories 0

dtype: int64

We will move forward with our analysis and decide if we want to predict those values to fill them with non-zero values.

|  | **activity\_calories** | **distance** | **floors** | **mins\_active\_high** | **mins\_active\_light** | **mins\_active\_med** | **mins\_sedentary** | **steps** | **total\_calories** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **count** | 225.000000 | 225.000000 | 225.000000 | 225.000000 | 225.000000 | 225.000000 | 225.000000 | 225.000000 | 225.000000 |
| **mean** | 1484.546667 | 4.046800 | 12.591111 | 33.871111 | 188.235556 | 14.920000 | 809.520000 | 8536.186667 | 3401.911111 |
| **std** | 747.795790 | 2.506253 | 25.844174 | 37.195986 | 75.170001 | 15.363535 | 243.232639 | 5255.746608 | 632.426578 |
| **min** | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 28.000000 | 0.000000 | 799.000000 |
| **25%** | 1062.000000 | 2.220000 | 2.000000 | 2.000000 | 158.000000 | 3.000000 | 654.000000 | 4714.000000 | 3069.000000 |
| **50%** | **1481.000000** | **3.680000** | **7.000000** | **15.000000** | **203.000000** | **11.000000** | **750.000000** | **7806.000000** | **3385.000000** |
| **75%** | 1962.000000 | 5.900000 | 13.000000 | 63.000000 | 236.000000 | 21.000000 | 877.000000 | 12380.000000 | 3765.000000 |
| **max** | 3596.000000 | 12.000000 | 195.000000 | 154.000000 | 349.000000 | 78.000000 | 1440.000000 | 25512.000000 | 5199.000000 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **activity\_calories** | **distance** | **floors** | **mins\_active\_high** | **mins\_active\_light** | **mins\_active\_med** | **mins\_sedentary** | **steps** | **total\_calories** |
| **1481.000000** | **3.680000** | **7.000000** | **15.000000** | **203.000000** | **11.000000** | **750.000000** | **7806.000000** | **3385.000000** |

Taking the median values for each feature, we can see that this person’s median daily distance travelled is 3.68 miles per day and the person median climb per day is 7 floors. The total activity adds up to only 16.3 hours, so the person must only be wearing the Fitbit while awake. That leaves 7.68 hours for sleep if he or she takes it off right before going to sleep and puts it on right after waking up. The person’s median sedentary time is 12.5 hours per day. The majority of activity is the 3.38 hours of active-light activity. So likely the person has a desk job, but works out, maybe also walking into and out of work and taking the stairs. Could work on the third floor of the building or the floor climbs could be done at the gym or at home.