insurance_costs

September 14, 2021

1 Medical Insurance Costs in the U.S.

1.1 Introduction

In this project we'll be analyzing a subset of insurance-oriented user data. This set includes information on customer age, sex, region, and annual medical insurance costs. Our goal will be to digest the information in this dataset into a more easily parsable form and analyze a few of its attributes through a helper class. Throughout the analysis we will be discussing the benefits and potential use of this data, and how it can be applied towards specific aims.

1.1.1 Imports

We only need to import python's innate csv functionality.

[6]:
```
import csv
```

After inspection of **insurance.csv** and ensuring that there are no hiccups - like missing data - we declare lists for each attribute we will be importing.

[7]:
```
ages = []
sexes = []
bmis = []
num_children = []
smoker_statuses = []
regions = []
insurance_charges = []
```

The file **insurance.csv** contains the columns: - Patient Age, continuous variable - Patient Sex, categorical variable - Patient BMI, continuous variable - Patient Number of Children, continuous variable - Patient Smoking Status, categorical variable - Patient U.S. Geographical Region, categorical variable - Patient Yearly Medical Insurance Cost, continuous variable

Using our declarations from earlier and a helper function, we'll be holding each column within a list.

[8]:
```
def load_list_data(data_list, csv_file, column):
    # opening the csv
    with open(csv_file) as csv_data:
        # reading data into a dictionary
        csv_dict = csv.DictReader(csv_data)
        # appending each row in the dict to a list
        for row in csv_dict:
            data_list.append(row[column])
        return data_list
```

```
[9]: load_list_data(ages, 'insurance.csv', 'age')
     load_list_data(sexes, 'insurance.csv', 'sex')
     load_list_data(bmis, 'insurance.csv', 'bmi')
     load_list_data(num_children, 'insurance.csv', 'children')
     load_list_data(smoker_statuses, 'insurance.csv', 'smoker')
     load_list_data(regions, 'insurance.csv', 'region')
     load_list_data(insurance_charges, 'insurance.csv', 'charges')
```

```
[9]: ['16884.924',
      '1725.5523',
      '4449.462',
      '21984.47061',
      '3866.8552',
      '3756.6216',
      '8240.5896',
      '7281.5056',
      '6406.4107',
      '28923.13692',
      '2721.3208',
      '27808.7251',
      '1826.843',
      '11090.7178',
      '39611.7577',
      '1837.237',
      '10797.3362',
      '2395.17155',
      '10602.385',
      '36837.467',
      '13228.84695',
      '4149.736',
      '1137.011',
      '37701.8768',
      '6203.90175',
      '14001.1338',
      '14451.83515',
      '12268.63225',
      '2775.19215',
      '38711',
      '35585.576',
      '2198.18985',
      '4687.797',
      '13770.0979',
      '51194.55914',
      '1625.43375',
      '15612.19335',
      '2302.3',
      '39774.2763',
      '48173.361',
      '3046.062',
      '4949.7587',
      '6272.4772',
      '6313.759',
      '6079.6715',
      '20630.28351',
      '3393.35635',
      '3556.9223',
      '12629.8967',
      '38709.176',
      '2211.13075',
      '3579.8287',
      '23568.272',
      '37742.5757',
      '8059.6791',
      '47496.49445',
      '13607.36875',
      '34303.1672',
      '23244.7902',
      '5989.52365',
      '8606.2174',
      '4504.6624',
      '30166.61817',
      '4133.64165',
      '14711.7438',
      '1743.214',
      '14235.072',
      '6389.37785',
      '5920.1041',
      '17663.1442',
      '16577.7795',
      '6799.458',
      '11741.726',
      '11946.6259',
      '7726.854',
      '11356.6609',
      '3947.4131',
      '1532.4697',
      '2755.02095',
      '6571.02435',
      '4441.21315',
      '7935.29115',
      '37165.1638',
      '11033.6617',
      '39836.519',
      '21098.55405',
      '43578.9394',
      '11073.176',
      '8026.6666',
      '11082.5772',
      '2026.9741',
      '10942.13205',
      '30184.9367',
      '5729.0053',
      '47291.055',
      '3766.8838',
      '12105.32',
      '10226.2842',
      '22412.6485',
      '15820.699',
      '6186.127',
      '3645.0894',
      '21344.8467',
      '30942.1918',
      '5003.853',
      '17560.37975',
      '2331.519',
      '3877.30425',
      '2867.1196',
      '47055.5321',
      '10825.2537',
      '11881.358',
      '4646.759',
      '2404.7338',
      '11488.31695',
      '30259.99556',
      '11381.3254',
      '19107.7796',
      '8601.3293',
      '6686.4313',
      '7740.337',
      '1705.6245',
      '2257.47525',
      '39556.4945',
      '10115.00885',
      '3385.39915',
      '17081.08',
      '9634.538',
      '32734.1863',
      '6082.405',
      '12815.44495',
      '13616.3586',
      '11163.568',
      '1632.56445',
      '2457.21115',
      '2155.6815',
      '1261.442',
      '2045.68525',
      '27322.73386',
      '2166.732',
      '27375.90478',
      '3490.5491',
      '18972.495',
      '18157.876',
      '20745.9891',
      '5138.2567',
      '40720.55105',
      '9877.6077',
      '10959.6947',
      '1842.519',
      '5125.2157',
      '7789.635',
      '6334.34355',
      '19964.7463',
      '7077.1894',
      '6948.7008',
      '21223.6758',
      '15518.18025',
      '36950.2567',
      '19749.38338',
      '21348.706',
      '36149.4835',
      '10450.552',
      '5152.134',
      '5028.1466',
      '10407.08585',
      '4830.63',
      '6128.79745',
      '2719.27975',
      '4827.90495',
      '13405.3903',
      '8116.68',
      '1694.7964',
      '5246.047',
      '2855.43755',
      '48824.45',
      '6455.86265',
      '10436.096',
      '8823.279',
      '8538.28845',
      '11735.87905',
      '1631.8212',
      '4005.4225',
      '7419.4779',
      '7731.4271',
      '43753.33705',
      '3981.9768',
      '5325.651',
      '6775.961',
      '4922.9159',
      '12557.6053',
      '4883.866',
      '2137.6536',
      '12044.342',
      '1137.4697',
      '1639.5631',
      '5649.715',
      '8516.829',
      '9644.2525',
      '14901.5167',
      '2130.6759',
      '8871.1517',
      '13012.20865',
      '37133.8982',
      '7147.105',
      '4337.7352',
      '11743.299',
      '20984.0936',
      '13880.949',
      '6610.1097',
      '1980.07',
      '8162.71625',
      '3537.703',
      '5002.7827',
      '8520.026',
      '7371.772',
      '10355.641',
      '2483.736',
      '3392.9768',
      '25081.76784',
      '5012.471',
      '10564.8845',
      '5253.524',
      '34779.615',
      '19515.5416',
      '11987.1682',
      '2689.4954',
      '24227.33724',
      '7358.17565',
      '9225.2564',
      '7443.64305',
      '14001.2867',
      '1727.785',
      '12333.828',
      '6710.1919',
      '19444.2658',
      '1615.7667',
      '4463.2051',
      '17352.6803',
      '7152.6714',
      '38511.6283',
      '5354.07465',
      '35160.13457',
      '7196.867',
      '29523.1656',
      '24476.47851',
      '12648.7034',
      '1986.9334',
      '1832.094',
      '4040.55825',
      '12829.4551',
      '47305.305',
      '44260.7499',
      '4260.744',
      '41097.16175',
      '13047.33235',
      '43921.1837',
      '5400.9805',
      '11520.09985',
      '33750.2918',
      '11837.16',
      '17085.2676',
      '24869.8368',
      '36219.40545',
      '20462.99766',
      '46151.1245',
      '17179.522',
      '14590.63205',
      '7441.053',
      '9282.4806',
      '1719.4363',
      '42856.838',
      '7265.7025',
      '9617.66245',
      '2523.1695',
      '9715.841',
      '2803.69785',
      '2150.469',
      '12928.7911',
      '9855.1314',
      '22331.5668',
      '48549.17835',
      '4237.12655',
      '11879.10405',
      '9625.92',
      '7742.1098',
      '9432.9253',
      '14256.1928',
      '47896.79135',
      '25992.82104',
      '3172.018',
      '20277.80751',
      '42112.2356',
      '2156.7518',
      '3906.127',
      '1704.5681',
      '16297.846',
      '21978.6769',
      '38746.3551',
      '9249.4952',
      '6746.7425',
      '24873.3849',
      '12265.5069',
      '4349.462',
      '12646.207',
      '19442.3535',
      '20177.67113',
      '4151.0287',
      '11944.59435',
      '7749.1564',
      '8444.474',
      '1737.376',
      '42124.5153',
      '8124.4084',
      '34838.873',
      '9722.7695',
      '8835.26495',
      '10435.06525',
      '7421.19455',
      '4667.60765',
      '4894.7533',
      '24671.66334',
      '35491.64',
      '11566.30055',
      '2866.091',
      '6600.20595',
      '3561.8889',
      '42760.5022',
      '47928.03',
      '9144.565',
      '48517.56315',
      '24393.6224',
      '13429.0354',
      '11658.37915',
      '19144.57652',
      '13822.803',
      '12142.5786',
      '13937.6665',
      '41919.097',
      '8232.6388',
      '18955.22017',
      '13352.0998',
      '13217.0945',
      '13981.85035',
      '10977.2063',
      '6184.2994',
      '4889.9995',
      '8334.45755',
      '5478.0368',
      '1635.73365',
      '11830.6072',
      '8932.084',
      '3554.203',
      '12404.8791',
      '14133.03775',
      '24603.04837',
      '8944.1151',
      '9620.3307',
      '1837.2819',
      '1607.5101',
      '10043.249',
      '4751.07',
      '13844.506',
      '2597.779',
      '3180.5101',
      '9778.3472',
      '13430.265',
      '8017.06115',
      '8116.26885',
      '3481.868',
      '13415.0381',
      '12029.2867',
      '7639.41745',
      '36085.219',
      '1391.5287',
      '18033.9679',
      '21659.9301',
      '38126.2465',
      '16455.70785',
      '27000.98473',
      '15006.57945',
      '42303.69215',
      '20781.48892',
      '5846.9176',
      '8302.53565',
      '1261.859',
      '11856.4115',
      '30284.64294',
      '3176.8159',
      '4618.0799',
      '10736.87075',
      '2138.0707',
      '8964.06055',
      '9290.1395',
      '9411.005',
      '7526.70645',
      '8522.003',
      '16586.49771',
      '14988.432',
      '1631.6683',
      '9264.797',
      '8083.9198',
      '14692.66935',
      '10269.46',
      '3260.199',
      '11396.9002',
      '4185.0979',
      '8539.671',
      '6652.5288',
      '4074.4537',
      '1621.3402',
      '19594.80965',
      '14455.64405',
      '5080.096',
      '2134.9015',
      '7345.7266',
      '9140.951',
      '18608.262',
      '14418.2804',
      '28950.4692',
      '46889.2612',
      '46599.1084',
      '39125.33225',
      '2727.3951',
      '8968.33',
      '9788.8659',
      '6555.07035',
      '7323.734819',
      '3167.45585',
      '18804.7524',
      '23082.95533',
      '4906.40965',
      '5969.723',
      '12638.195',
      '4243.59005',
      '13919.8229',
      '2254.7967',
      '5926.846',
      '12592.5345',
      '2897.3235',
      '4738.2682',
      '37079.372',
      '1149.3959',
      '28287.89766',
      '26109.32905',
      '7345.084',
      '12730.9996',
      '11454.0215',
      '5910.944',
      '4762.329',
      '7512.267',
      '4032.2407',
      '1969.614',
      '1769.53165',
      '4686.3887',
      '21797.0004',
      '11881.9696',
      '11840.77505',
      '10601.412',
      '7682.67',
      '10381.4787',
      '22144.032',
      '15230.32405',
      '11165.41765',
      '1632.03625',
      '19521.9682',
      '13224.693',
      '12643.3778',
      '23288.9284',
      '2201.0971',
      '2497.0383',
      '2203.47185',
      '1744.465',
      '20878.78443',
      '25382.297',
      '28868.6639',
      '35147.52848',
      '2534.39375',
      '1534.3045',
      '1824.2854',
      '15555.18875',
      '9304.7019',
      '1622.1885',
      '9880.068',
      '9563.029',
      '4347.02335',
      '12475.3513',
      '1253.936',
      '48885.13561',
      '10461.9794',
      '1748.774',
      '24513.09126',
      '2196.4732',
      '12574.049',
      '17942.106',
      '1967.0227',
      '4931.647',
      '8027.968',
      '8211.1002',
      '13470.86',
      '36197.699',
      '6837.3687',
      '22218.1149',
      '32548.3405',
      '5974.3847',
      '6796.86325',
      '2643.2685',
      '3077.0955',
      '3044.2133',
      '11455.28',
      '11763.0009',
      '2498.4144',
      '9361.3268',
      '1256.299',
      '21082.16',
      '11362.755',
      '27724.28875',
      '8413.46305',
      '5240.765',
      '3857.75925',
      '25656.57526',
      '3994.1778',
      '9866.30485',
      '5397.6167',
      '38245.59327',
      '11482.63485',
      '24059.68019',
      '9861.025',
      '8342.90875',
      '1708.0014',
      '48675.5177',
      '14043.4767',
      '12925.886',
      '19214.70553',
      '13831.1152',
      '6067.12675',
      '5972.378',
      '8825.086',
      '8233.0975',
      '27346.04207',
      '6196.448',
      '3056.3881',
      '13887.204',
      '63770.42801',
      '10231.4999',
      '23807.2406',
      '3268.84665',
      '11538.421',
      '3213.62205',
      '45863.205',
      '13390.559',
      '3972.9247',
      '12957.118',
      '11187.6567',
      '17878.90068',
      '3847.674',
      '8334.5896',
      '3935.1799',
      '39983.42595',
      '1646.4297',
      '9193.8385',
      '10923.9332',
      '2494.022',
      '9058.7303',
      '2801.2588',
      '2128.43105',
      '6373.55735',
      '7256.7231',
      '11552.904',
      '45702.02235',
      '3761.292',
      '2219.4451',
      '4753.6368',
      '31620.00106',
      '13224.05705',
      '12222.8983',
      '1664.9996',
      '58571.07448',
      '9724.53',
      '3206.49135',
      '12913.9924',
      '1639.5631',
      '6356.2707',
      '17626.23951',
      '1242.816',
      '4779.6023',
      '3861.20965',
      '43943.8761',
      '13635.6379',
      '5976.8311',
      '11842.442',
      '8428.0693',
      '2566.4707',
      '15359.1045',
      '5709.1644',
      '8823.98575',
      '7640.3092',
      '5594.8455',
      '7441.501',
      '33471.97189',
      '1633.0444',
      '9174.13565',
      '11070.535',
      '16085.1275',
      '17468.9839',
      '9283.562',
      '3558.62025',
      '25678.77845',
      '4435.0942',
      '39241.442',
      '8547.6913',
      '6571.544',
      '2207.69745',
      '6753.038',
      '1880.07',
      '42969.8527',
      '11658.11505',
      '23306.547',
      '34439.8559',
      '10713.644',
      '3659.346',
      '40182.246',
      '9182.17',
      '34617.84065',
      '12129.61415',
      '3736.4647',
      '6748.5912',
      '11326.71487',
      '11365.952',
      '42983.4585',
      '10085.846',
      '1977.815',
      '3366.6697',
      '7173.35995',
      '9391.346',
      '14410.9321',
      '2709.1119',
      '24915.04626',
      '20149.3229',
      '12949.1554',
      '6666.243',
      '32787.45859',
      '13143.86485',
      '4466.6214',
      '18806.14547',
      '10141.1362',
      '6123.5688',
      '8252.2843',
      '1712.227',
      '12430.95335',
      '9800.8882',
      '10579.711',
      '8280.6227',
      '8527.532',
      '12244.531',
      '24667.419',
      '3410.324',
      '4058.71245',
      '26392.26029',
      '14394.39815',
      '6435.6237',
      '22192.43711',
      '5148.5526',
      '1136.3994',
      '27037.9141',
      '42560.4304',
      '8703.456',
      '40003.33225',
      '45710.20785',
      '6500.2359',
      '4837.5823',
      '3943.5954',
      '4399.731',
      '6185.3208',
      '46200.9851',
      '7222.78625',
      '12485.8009',
      '46130.5265',
      '12363.547',
      '10156.7832',
      '2585.269',
      '1242.26',
      '40103.89',
      '9863.4718',
      '4766.022',
      '11244.3769',
      '7729.64575',
      '5438.7491',
      '26236.57997',
      '34806.4677',
      '2104.1134',
      '8068.185',
      '2362.22905',
      '2352.96845',
      '3577.999',
      '3201.24515',
      '29186.48236',
      '40273.6455',
      '10976.24575',
      '3500.6123',
      '2020.5523',
      '9541.69555',
      '9504.3103',
      '5385.3379',
      '8930.93455',
      '5375.038',
      '44400.4064',
      '10264.4421',
      '6113.23105',
      '5469.0066',
      '1727.54',
      '10107.2206',
      '8310.83915',
      '1984.4533',
      '2457.502',
      '12146.971',
      '9566.9909',
      '13112.6048',
      '10848.1343',
      '12231.6136',
      '9875.6804',
      '11264.541',
      '12979.358',
      '1263.249',
      '10106.13425',
      '40932.4295',
      '6664.68595',
      '16657.71745',
      '2217.6012',
      '6781.3542',
      '19361.9988',
      '10065.413',
      '4234.927',
      '9447.25035',
      '14007.222',
      '9583.8933',
      '40419.0191',
      '3484.331',
      '36189.1017',
      '44585.45587',
      '8604.48365',
      '18246.4955',
      '43254.41795',
      '3757.8448',
      '8827.2099',
      '9910.35985',
      '11737.84884',
      '1627.28245',
      '8556.907',
      '3062.50825',
      '19539.243',
      '1906.35825',
      '14210.53595',
      '11833.7823',
      '17128.42608',
      '5031.26955',
      '7985.815',
      '23065.4207',
      '5428.7277',
      '36307.7983',
      '3925.7582',
      '2416.955',
      '19040.876',
      '3070.8087',
      '9095.06825',
      '11842.62375',
      '8062.764',
      '7050.642',
      '14319.031',
      '6933.24225',
      '27941.28758',
      '11150.78',
      '12797.20962',
      '17748.5062',
      '7261.741',
      '10560.4917',
      '6986.697',
      '7448.40395',
      '5934.3798',
      '9869.8102',
      '18259.216',
      '1146.7966',
      '9386.1613',
      '24520.264',
      '4350.5144',
      '6414.178',
      '12741.16745',
      '1917.3184',
      '5209.57885',
      '13457.9608',
      '5662.225',
      '1252.407',
      '2731.9122',
      '21195.818',
      '7209.4918',
      '18310.742',
      '4266.1658',
      '4719.52405',
      '11848.141',
      '17904.52705',
      '7046.7222',
      '14313.8463',
      '2103.08',
      '38792.6856',
      '1815.8759',
      '7731.85785',
      '28476.73499',
      '2136.88225',
      '1131.5066',
      '3309.7926',
      '9414.92',
      '6360.9936',
      '11013.7119',
      '4428.88785',
      '5584.3057',
      '1877.9294',
      '2842.76075',
      '3597.596',
      '23401.30575',
      '55135.40209',
      '7445.918',
      '2680.9493',
      '1621.8827',
      '8219.2039',
      '12523.6048',
      '16069.08475',
      '43813.8661',
      '20773.62775',
      '39597.4072',
      '6117.4945',
      '13393.756',
      '5266.3656',
      '4719.73655',
      '11743.9341',
      '5377.4578',
      '7160.3303',
      '4402.233',
      '11657.7189',
      '6402.29135',
      '12622.1795',
      '1526.312',
      '12323.936',
      '36021.0112',
      '27533.9129',
      '10072.05505',
      '45008.9555',
      '9872.701',
      '2438.0552',
      '2974.126',
      '10601.63225',
      '37270.1512',
      '14119.62',
      '42111.6647',
      '11729.6795',
      '24106.91255',
      '1875.344',
      '40974.1649',
      '15817.9857',
      '18218.16139',
      '10965.446',
      '46113.511',
      '7151.092',
      '12269.68865',
      '5458.04645',
      '8782.469',
      '6600.361',
      '1141.4451',
      '11576.13',
      '13129.60345',
      '4391.652',
      '8457.818',
      '3392.3652',
      '5966.8874',
      '6849.026',
      '8891.1395',
      '2690.1138',
      '26140.3603',
      '6653.7886',
      '6282.235',
      '6311.952',
      '3443.064',
      '2789.0574',
      '2585.85065',
      '46255.1125',
      '4877.98105',
      '19719.6947',
      '27218.43725',
      '5272.1758',
      '1682.597',
      '11945.1327',
      '29330.98315',
      '7243.8136',
      '10422.91665',
      '44202.6536',
      '13555.0049',
      '13063.883',
      '19798.05455',
      '2221.56445',
      '1634.5734',
      '2117.33885',
      '8688.85885',
      '48673.5588',
      '4661.28635',
      '8125.7845',
      '12644.589',
      '4564.19145',
      '4846.92015',
      '7633.7206',
      '15170.069',
      '17496.306',
      '2639.0429',
      '33732.6867',
      '14382.70905',
      '7626.993',
      '5257.50795',
      '2473.3341',
      '21774.32215',
      '35069.37452',
      '13041.921',
      '5245.2269',
      '13451.122',
      '13462.52',
      '5488.262',
      '4320.41085',
      '6250.435',
      '25333.33284',
      '2913.569',
      '12032.326',
      '13470.8044',
      '6289.7549',
      '2927.0647',
      '6238.298',
      '10096.97',
      '7348.142',
      '4673.3922',
      '12233.828',
      '32108.66282',
      '8965.79575',
      '2304.0022',
      '9487.6442',
      '1121.8739',
      '9549.5651',
      '2217.46915',
      '1628.4709',
      '12982.8747',
      '11674.13',
      '7160.094',
      '39047.285',
      '6358.77645',
      '19933.458',
      '11534.87265',
      '47462.894',
      '4527.18295',
      '38998.546',
      '20009.63365',
      '3875.7341',
      '41999.52',
      '12609.88702',
      '41034.2214',
      '28468.91901',
      '2730.10785',
      '3353.284',
      '14474.675',
      '9500.57305',
      '26467.09737',
      '4746.344',
      '23967.38305',
      '7518.02535',
      '3279.86855',
      '8596.8278',
      '10702.6424',
      '4992.3764',
      '2527.81865',
      '1759.338',
      '2322.6218',
      '16138.76205',
      '7804.1605',
      '2902.9065',
      '9704.66805',
      '4889.0368',
      '25517.11363',
      '4500.33925',
      '19199.944',
      '16796.41194',
      '4915.05985',
      '7624.63',
      '8410.04685',
      '28340.18885',
      '4518.82625',
      '14571.8908',
      '3378.91',
      '7144.86265',
      '10118.424',
      '5484.4673',
      '16420.49455',
      '7986.47525',
      '7418.522',
      '13887.9685',
      '6551.7501',
      '5267.81815',
      …]
```

Since the data has been organized into lists we can perform our analysis. Although we can focus on a number of aspects, in this project we will be investigating: - The average patient age - The sex patient sex distribution - The geographical location - The average yearly medical costs - An easily searchable patient dictionary

1.2 Data Analysis

To accomplish this, we will create a class with methods that will help us parse through the data.

```
[10]: class PtInfo:
          # bringing in the values from the data and initializing them
          def __init__(self, patients_ages, patients_sexes, patients_bmis,␣
       ↪patients_num_children,
                       patients_smoker_statuses, patients_regions, patients_charges):
              self.patients_ages = patients_ages
              self.patients_sexes = patients_sexes
              self.patients_bmis = patients_bmis
              self.patients_num_children = patients_num_children
              self.patients_smoker_statuses = patients_smoker_statuses
              self.patients_regions = patients_regions
```

```
        self.patients_charges = patients_charges

    # calculating the average age of the patients
    def analyze_ages(self):
        # declaring the total_age var and initializing it at 0 - this is our␣
↪counts
        total_age = 0
        # iterating through all the patient ages
        for age in self.patients_ages:
            # summing them
            total_age += int(age)
        # returning the average total age
        return (f'Average Patient Age: {str(round(total_age/len(self.
↪patients_ages), 2))} years')

    # calculating the male / female distribution
    def analyze_sexes(self):
        # creating our count vars
        females = 0
        males = 0
        # iterating and counting
        for sex in self.patients_sexes:
            # if female add to female variable
            if sex == 'female':
                females += 1
            # if male add to male variable
            elif sex == 'male':
                males += 1
        # printing the number of males and females in the set
        print(f'Count for female: {females}')
        print(f'Count for male: {males}')

    # finding each unique region the patients are from
    def unique_regions(self):
        # creating an empty list to hold each new region
        unique_regions = []
        # iterating through the regions
        for region in self.patients_regions:
            # if not in our region list, adding it to the list
            if region not in unique_regions:
                unique_regions.append(region)
        # returning the list
        return unique_regions

    # determing average yearly medical insurance cost
    def average_charges(self):
        # initializing total_charges
        total_charges = 0
        # iterating through each charge and adding it to the total
        for charge in self.patients_charges:
            total_charges += float(charge)
        # returning the average charges, rounded to the nearest hundredths place
        return (f'Average Yearly Medical Insurance Charges:␣
↪{str(round(total_charges/len(self.patients_charges), 2))} dollars.')

    # creating a dictionary with all the patient info
    def create_dictionary(self):
        # initializing the dictionary
        self.patients_dictionary = {}
        # iterating through each list and adding it to the dict
        self.patients_dictionary["age"] = [int(age) for age in self.
↪patients_ages]
        self.patients_dictionary["sex"] = self.patients_sexes
        self.patients_dictionary["bmi"] = self.patients_bmis
        self.patients_dictionary["children"] = self.patients_num_children
        self.patients_dictionary["smoker"] = self.patients_smoker_statuses
        self.patients_dictionary["regions"] = self.patients_regions
        self.patients_dictionary["charges"] = self.patients_charges
        return self.patients_dictionary
```

Now we instantiate can use our new class and utilize the methods we built.

[11]:
```
patient_info = PtInfo(ages, sexes, bmis, num_children, smoker_statuses,␣
↪regions, insurance_charges)
```

The average patient age is 39 years. This is important for a wide range of use cases, such as what demographic to focus a hypothetical new advertisement campaign on. This could also be of use when determining risk tolerance.

Further analysis could focus on potential data skew of the population.

[12]:
```
print(patient_info.analyze_ages())
```

```
Average Patient Age: 39.21 years
```

Here we are checking the balance of males in comparison to females in the dataset. Surprisingly they are near equal, with males outnumbering females by only 6 persons.

[13]:
```
print(patient_info.analyze_sexes())
```

```
Count for female: 662
Count for male: 676
None
```

There are four regions of the U.S. represented in this set: Southwest, Southeast, Northwest, and Northeast. Further analysis could study breakdown of patient health factors and their prevalence in each region.

[14]:
```
print(patient_info.unique_regions())
```

```
['southwest', 'southeast', 'northwest', 'northeast']
```

From this sampling, the average yearly cost for medical insurance is $13,270.42 per individual. We could also further compare patient health identifiers and their impact on insurance cost. One could imagine that BMI or age could correlate in a significant way with the yearly use of medical insurance, and thus its annual cost.

[15]:
```
print(patient_info.average_charges())
```

[15]:
```
'Average Yearly Medical Insurance Charges: 13270.42 dollars.'
```

Finally, we will load our patient data into an easily searchable dictionary to make further analyses easier.

[ ]:
```
ptdict = patient_info.create_dictionary()
for key, value in ptdict.items():
    print(key, value)
```

1.3 Conclusion

Using basic python, we utilized our subset of insurance data to produce population, region, and cost descriptions. We discussed what aims these features could be used for and identified areas of further investigation. Finally, we transformed our data into a searchable dictionary to make later analysis easier.

Data Sources: The insurance data (`insurance.csv`) was provided by Codecademy.com.