

# assignment\_03(02)\_PothineniKalyan

PothineniKalyan

2023-03-31

```
## Load the ggplot2 package
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
theme_set(theme_minimal())
```

```
## Set the working directory to the root of your DSC 520 directory
setwd("C:/Users/kpothine/OneDrive - Waste Management/Documents/NDO_GIT/dsc520")
```

```
## Load the `data/acs-14-1yr-s0201.csv` to
us_census_df <- read.csv("data/acs-14-1yr-s0201.csv")
```

```
## List the data frame
list(us_census_df)
```

```
## [[1]]
##           Id   Id2 Geography PopGroupID
## 1 0500000US01073 1073 Jefferson County, Alabama      1
## 2 0500000US04013 4013 Maricopa County, Arizona      1
## 3 0500000US04019 4019 Pima County, Arizona      1
## 4 0500000US06001 6001 Alameda County, California      1
## 5 0500000US06013 6013 Contra Costa County, California      1
## 6 0500000US06019 6019 Fresno County, California      1
## 7 0500000US06029 6029 Kern County, California      1
## 8 0500000US06037 6037 Los Angeles County, California      1
## 9 0500000US06059 6059 Orange County, California      1
## 10 0500000US06065 6065 Riverside County, California      1
## 11 0500000US06067 6067 Sacramento County, California      1
## 12 0500000US06071 6071 San Bernardino County, California      1
## 13 0500000US06073 6073 San Diego County, California      1
## 14 0500000US06075 6075 San Francisco County, California      1
## 15 0500000US06077 6077 San Joaquin County, California      1
## 16 0500000US06081 6081 San Mateo County, California      1
## 17 0500000US06085 6085 Santa Clara County, California      1
## 18 0500000US06097 6097 Sonoma County, California      1
## 19 0500000US06099 6099 Stanislaus County, California      1
## 20 0500000US06111 6111 Ventura County, California      1
## 21 0500000US08005 8005 Arapahoe County, Colorado      1
## 22 0500000US08031 8031 Denver County, Colorado      1
```

## 23	0500000US08041	8041	El Paso County, Colorado	1
## 24	0500000US08059	8059	Jefferson County, Colorado	1
## 25	0500000US09001	9001	Fairfield County, Connecticut	1
## 26	0500000US09003	9003	Hartford County, Connecticut	1
## 27	0500000US09009	9009	New Haven County, Connecticut	1
## 28	0500000US10003	10003	New Castle County, Delaware	1
## 29	0500000US11001	11001	District of Columbia, District of Columbia	1
## 30	0500000US12009	12009	Brevard County, Florida	1
## 31	0500000US12011	12011	Broward County, Florida	1
## 32	0500000US12031	12031	Duval County, Florida	1
## 33	0500000US12057	12057	Hillsborough County, Florida	1
## 34	0500000US12071	12071	Lee County, Florida	1
## 35	0500000US12086	12086	Miami-Dade County, Florida	1
## 36	0500000US12095	12095	Orange County, Florida	1
## 37	0500000US12099	12099	Palm Beach County, Florida	1
## 38	0500000US12103	12103	Pinellas County, Florida	1
## 39	0500000US12105	12105	Polk County, Florida	1
## 40	0500000US12127	12127	Volusia County, Florida	1
## 41	0500000US13067	13067	Cobb County, Georgia	1
## 42	0500000US13089	13089	DeKalb County, Georgia	1
## 43	0500000US13121	13121	Fulton County, Georgia	1
## 44	0500000US13135	13135	Gwinnett County, Georgia	1
## 45	0500000US15003	15003	Honolulu County, Hawaii	1
## 46	0500000US17031	17031	Cook County, Illinois	1
## 47	0500000US17043	17043	DuPage County, Illinois	1
## 48	0500000US17089	17089	Kane County, Illinois	1
## 49	0500000US17097	17097	Lake County, Illinois	1
## 50	0500000US17197	17197	Will County, Illinois	1
## 51	0500000US18097	18097	Marion County, Indiana	1
## 52	0500000US20091	20091	Johnson County, Kansas	1
## 53	0500000US20173	20173	Sedgwick County, Kansas	1
## 54	0500000US21111	21111	Jefferson County, Kentucky	1
## 55	0500000US24003	24003	Anne Arundel County, Maryland	1
## 56	0500000US24005	24005	Baltimore County, Maryland	1
## 57	0500000US24031	24031	Montgomery County, Maryland	1
## 58	0500000US24033	24033	Prince George's County, Maryland	1
## 59	0500000US24510	24510	Baltimore city, Maryland	1
## 60	0500000US25005	25005	Bristol County, Massachusetts	1
## 61	0500000US25009	25009	Essex County, Massachusetts	1
## 62	0500000US25017	25017	Middlesex County, Massachusetts	1
## 63	0500000US25021	25021	Norfolk County, Massachusetts	1
## 64	0500000US25023	25023	Plymouth County, Massachusetts	1
## 65	0500000US25025	25025	Suffolk County, Massachusetts	1
## 66	0500000US25027	25027	Worcester County, Massachusetts	1
## 67	0500000US26081	26081	Kent County, Michigan	1
## 68	0500000US26099	26099	Macomb County, Michigan	1
## 69	0500000US26125	26125	Oakland County, Michigan	1
## 70	0500000US26163	26163	Wayne County, Michigan	1
## 71	0500000US27053	27053	Hennepin County, Minnesota	1
## 72	0500000US27123	27123	Ramsey County, Minnesota	1
## 73	0500000US29095	29095	Jackson County, Missouri	1
## 74	0500000US29189	29189	St. Louis County, Missouri	1
## 75	0500000US31055	31055	Douglas County, Nebraska	1
## 76	0500000US32003	32003	Clark County, Nevada	1

## 77	0500000US34003	34003	Bergen County, New Jersey	1
## 78	0500000US34007	34007	Camden County, New Jersey	1
## 79	0500000US34013	34013	Essex County, New Jersey	1
## 80	0500000US34017	34017	Hudson County, New Jersey	1
## 81	0500000US34023	34023	Middlesex County, New Jersey	1
## 82	0500000US34025	34025	Monmouth County, New Jersey	1
## 83	0500000US34029	34029	Ocean County, New Jersey	1
## 84	0500000US34031	34031	Passaic County, New Jersey	1
## 85	0500000US34039	34039	Union County, New Jersey	1
## 86	0500000US35001	35001	Bernalillo County, New Mexico	1
## 87	0500000US36005	36005	Bronx County, New York	1
## 88	0500000US36029	36029	Erie County, New York	1
## 89	0500000US36047	36047	Kings County, New York	1
## 90	0500000US36055	36055	Monroe County, New York	1
## 91	0500000US36059	36059	Nassau County, New York	1
## 92	0500000US36061	36061	New York County, New York	1
## 93	0500000US36081	36081	Queens County, New York	1
## 94	0500000US36103	36103	Suffolk County, New York	1
## 95	0500000US36119	36119	Westchester County, New York	1
## 96	0500000US37081	37081	Guilford County, North Carolina	1
## 97	0500000US37119	37119	Mecklenburg County, North Carolina	1
## 98	0500000US37183	37183	Wake County, North Carolina	1
## 99	0500000US39035	39035	Cuyahoga County, Ohio	1
## 100	0500000US39049	39049	Franklin County, Ohio	1
## 101	0500000US39061	39061	Hamilton County, Ohio	1
## 102	0500000US39113	39113	Montgomery County, Ohio	1
## 103	0500000US39153	39153	Summit County, Ohio	1
## 104	0500000US40109	40109	Oklahoma County, Oklahoma	1
## 105	0500000US40143	40143	Tulsa County, Oklahoma	1
## 106	0500000US41051	41051	Multnomah County, Oregon	1
## 107	0500000US41067	41067	Washington County, Oregon	1
## 108	0500000US42003	42003	Allegheny County, Pennsylvania	1
## 109	0500000US42017	42017	Bucks County, Pennsylvania	1
## 110	0500000US42029	42029	Chester County, Pennsylvania	1
## 111	0500000US42045	42045	Delaware County, Pennsylvania	1
## 112	0500000US42071	42071	Lancaster County, Pennsylvania	1
## 113	0500000US42091	42091	Montgomery County, Pennsylvania	1
## 114	0500000US42101	42101	Philadelphia County, Pennsylvania	1
## 115	0500000US44007	44007	Providence County, Rhode Island	1
## 116	0500000US47037	47037	Davidson County, Tennessee	1
## 117	0500000US47157	47157	Shelby County, Tennessee	1
## 118	0500000US48029	48029	Bexar County, Texas	1
## 119	0500000US48085	48085	Collin County, Texas	1
## 120	0500000US48113	48113	Dallas County, Texas	1
## 121	0500000US48121	48121	Denton County, Texas	1
## 122	0500000US48141	48141	El Paso County, Texas	1
## 123	0500000US48157	48157	Fort Bend County, Texas	1
## 124	0500000US48201	48201	Harris County, Texas	1
## 125	0500000US48215	48215	Hidalgo County, Texas	1
## 126	0500000US48339	48339	Montgomery County, Texas	1
## 127	0500000US48439	48439	Tarrant County, Texas	1
## 128	0500000US48453	48453	Travis County, Texas	1
## 129	0500000US49035	49035	Salt Lake County, Utah	1
## 130	0500000US49049	49049	Utah County, Utah	1

## 131	0500000US51059	51059	Fairfax County, Virginia	1
## 132	0500000US53033	53033	King County, Washington	1
## 133	0500000US53053	53053	Pierce County, Washington	1
## 134	0500000US53061	53061	Snohomish County, Washington	1
## 135	0500000US55025	55025	Dane County, Wisconsin	1
## 136	0500000US55079	55079	Milwaukee County, Wisconsin	1
##	POPGRPUP.display.label	RacesReported	HSDegree	BachDegree
## 1	Total population	660793	89.1	30.5
## 2	Total population	4087191	86.8	30.2
## 3	Total population	1004516	88.0	30.8
## 4	Total population	1610921	86.9	42.8
## 5	Total population	1111339	88.8	39.7
## 6	Total population	965974	73.6	19.7
## 7	Total population	874589	74.5	15.4
## 8	Total population	10116705	77.5	30.3
## 9	Total population	3145515	84.6	38.0
## 10	Total population	2329271	80.6	20.7
## 11	Total population	1482026	86.8	28.9
## 12	Total population	2112619	78.6	18.9
## 13	Total population	3263431	86.6	37.1
## 14	Total population	852469	88.1	54.2
## 15	Total population	715597	77.6	18.3
## 16	Total population	758581	88.1	47.5
## 17	Total population	1894605	87.4	48.4
## 18	Total population	500292	87.6	34.8
## 19	Total population	531997	78.4	17.0
## 20	Total population	846178	83.6	31.6
## 21	Total population	618821	91.9	40.9
## 22	Total population	663862	85.5	44.3
## 23	Total population	663519	92.8	36.5
## 24	Total population	558503	94.1	42.0
## 25	Total population	945438	89.8	46.7
## 26	Total population	897985	89.3	36.8
## 27	Total population	861277	89.5	34.5
## 28	Total population	552778	90.1	35.8
## 29	Total population	658893	90.2	55.0
## 30	Total population	556885	91.6	27.2
## 31	Total population	1869235	88.4	30.5
## 32	Total population	897698	89.0	26.1
## 33	Total population	1316298	87.3	29.8
## 34	Total population	679513	86.3	26.5
## 35	Total population	2662874	80.9	26.6
## 36	Total population	1253001	87.9	31.4
## 37	Total population	1397710	87.7	33.0
## 38	Total population	938098	90.1	29.5
## 39	Total population	634638	84.9	19.7
## 40	Total population	507531	88.9	22.5
## 41	Total population	730981	90.3	43.7
## 42	Total population	722161	88.4	41.7
## 43	Total population	996319	91.3	49.2
## 44	Total population	877922	88.0	35.4
## 45	Total population	991788	91.8	32.6
## 46	Total population	5246456	85.5	36.2
## 47	Total population	932708	92.3	48.0

## 48	Total population	527306	82.9	32.6
## 49	Total population	705186	90.3	44.0
## 50	Total population	685419	90.7	33.1
## 51	Total population	934243	85.0	28.8
## 52	Total population	574272	95.5	52.8
## 53	Total population	508803	88.8	30.7
## 54	Total population	760026	88.5	31.6
## 55	Total population	560133	91.9	38.8
## 56	Total population	826925	90.4	37.2
## 57	Total population	1030447	90.9	58.5
## 58	Total population	904430	85.5	31.0
## 59	Total population	622793	84.4	30.0
## 60	Total population	554194	82.5	25.7
## 61	Total population	769091	89.1	38.9
## 62	Total population	1570315	92.3	52.3
## 63	Total population	692254	94.1	51.9
## 64	Total population	507022	92.2	34.1
## 65	Total population	767254	83.9	42.3
## 66	Total population	813475	90.1	34.6
## 67	Total population	629237	89.1	33.7
## 68	Total population	860112	89.3	23.9
## 69	Total population	1237868	93.6	44.8
## 70	Total population	1764804	84.9	22.1
## 71	Total population	1212064	93.2	47.3
## 72	Total population	532655	89.9	40.9
## 73	Total population	683191	90.0	29.5
## 74	Total population	1001876	93.2	42.8
## 75	Total population	543244	88.2	36.3
## 76	Total population	2069681	84.5	22.7
## 77	Total population	933572	91.5	46.2
## 78	Total population	511038	88.3	31.3
## 79	Total population	795723	85.5	32.7
## 80	Total population	669115	83.4	38.2
## 81	Total population	836297	89.1	41.0
## 82	Total population	629279	93.1	43.7
## 83	Total population	586301	91.7	28.6
## 84	Total population	508856	83.8	28.6
## 85	Total population	552939	86.2	33.0
## 86	Total population	675551	88.0	32.7
## 87	Total population	1438159	70.5	19.3
## 88	Total population	922835	90.6	31.3
## 89	Total population	2621793	80.0	34.3
## 90	Total population	749857	90.3	35.9
## 91	Total population	1358627	90.7	43.2
## 92	Total population	1636268	86.8	59.9
## 93	Total population	2321580	80.4	29.8
## 94	Total population	1502968	89.8	34.0
## 95	Total population	972634	87.4	47.1
## 96	Total population	512119	89.0	33.3
## 97	Total population	1012539	89.5	43.0
## 98	Total population	998691	92.4	49.2
## 99	Total population	1259828	88.1	31.0
## 100	Total population	1231393	90.0	38.0
## 101	Total population	806631	90.5	35.6

```
## 102      Total population      533116      89.7      25.7
## 103      Total population      541943      91.1      30.3
## 104      Total population      766215      86.8      30.6
## 105      Total population      629598      88.6      30.7
## 106      Total population      776712      91.1      41.6
## 107      Total population      562998      90.2      39.7
## 108      Total population     1231255      93.9      37.7
## 109      Total population      626685      93.9      37.7
## 110      Total population      512784      92.3      49.3
## 111      Total population      562960      91.5      36.3
## 112      Total population      533320      84.9      26.0
## 113      Total population      816857      93.7      47.3
## 114      Total population     1560297      82.6      26.0
## 115      Total population      631974      82.0      25.2
## 116      Total population      668347      86.7      37.3
## 117      Total population      938803      87.4      29.9
## 118      Total population     1855866      83.0      26.3
## 119      Total population      885241      93.7      50.0
## 120      Total population     2518638      77.6      29.1
## 121      Total population      753363      91.9      41.5
## 122      Total population      833487      75.8      21.1
## 123      Total population      685345      88.6      44.1
## 124      Total population     4441370      79.8      29.7
## 125      Total population      831073      62.2      17.9
## 126      Total population      518947      85.9      34.1
## 127      Total population     1945360      84.9      30.0
## 128      Total population     1151145      88.6      45.6
## 129      Total population     1091742      89.5      31.9
## 130      Total population      560974      93.7      37.5
## 131      Total population     1137538      91.5      60.3
## 132      Total population     2079967      92.3      48.6
## 133      Total population      831928      90.3      24.6
## 134      Total population      759583      92.0      29.1
## 135      Total population      516284      94.9      49.8
## 136      Total population      956406      86.9      29.5
```

```
## Structure of data frame
```

```
str(us_census_df)
```

```
## 'data.frame':   136 obs. of  8 variables:
## $ Id           : chr  "05000000US01073" "05000000US04013" "05000000US04019" "05000000US06001"
## $ Id2          : int   1073 4013 4019 6001 6013 6019 6029 6037 6059 6065 ...
## $ Geography    : chr   "Jefferson County, Alabama" "Maricopa County, Arizona" "Pima County,
## $ PopGroupID   : int    1 1 1 1 1 1 1 1 1 1 ...
## $ POPGROUP.display.label: chr  "Total population" "Total population" "Total population" "Total popu
## $ RacesReported : int   660793 4087191 1004516 1610921 1111339 965974 874589 10116705 314551
## $ HSDegree     : num    89.1 86.8 88 86.9 88.8 73.6 74.5 77.5 84.6 80.6 ...
## $ BachDegree   : num    30.5 30.2 30.8 42.8 39.7 19.7 15.4 30.3 38 20.7 ...
```

```
## Number of rows in the data frame
```

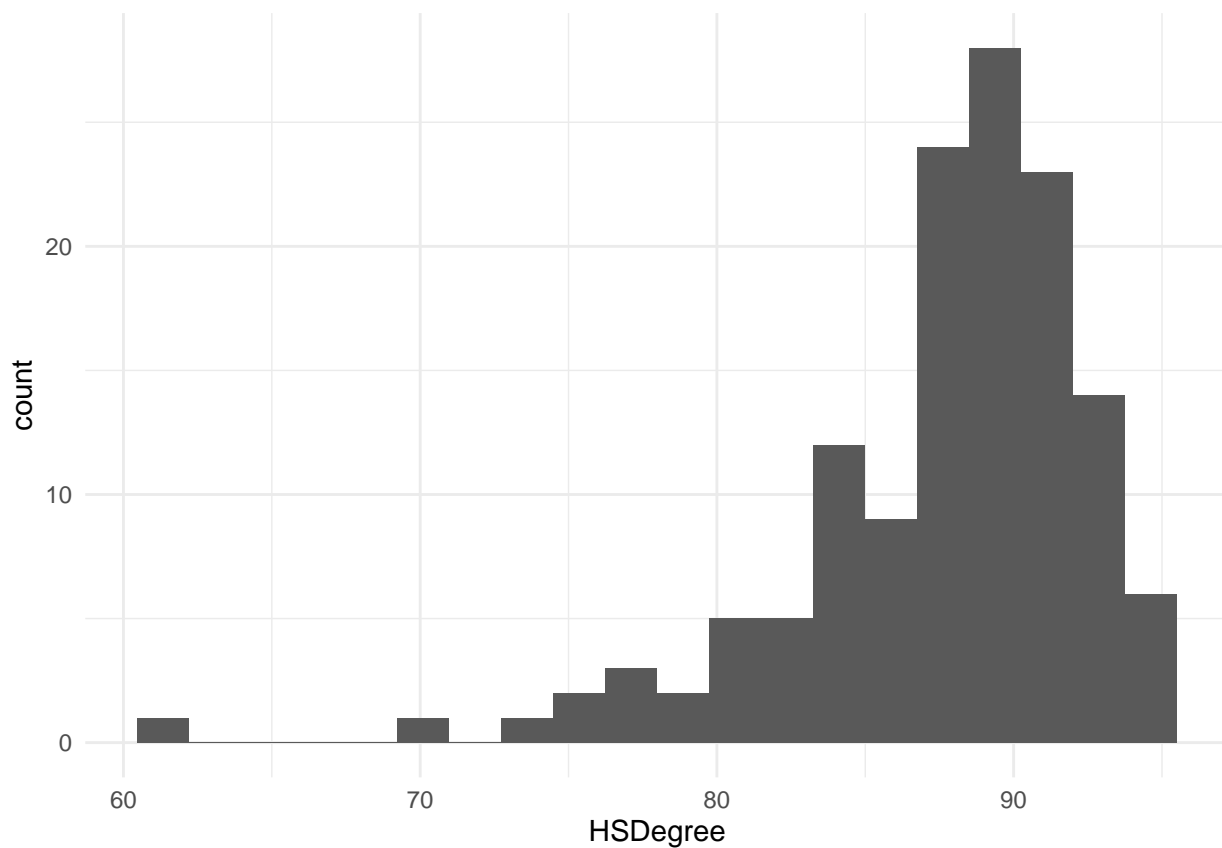
```
nrow(us_census_df)
```

```
## [1] 136
```

```
## Number of columns in the data frame
ncol(us_census_df)
```

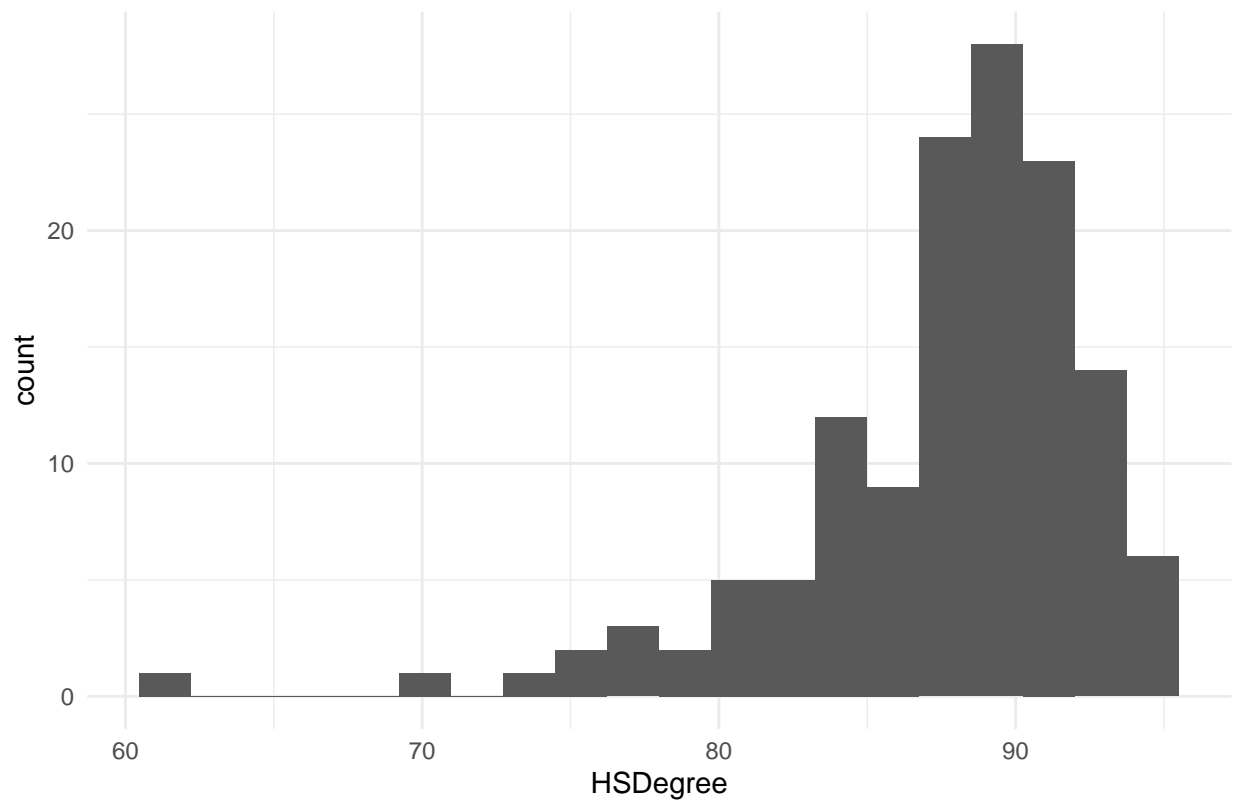
```
## [1] 8
```

```
## Histogram of the `HSDegree` variable using `geom_histogram()`
ggplot(us_census_df, aes(HSDegree)) + geom_histogram( bins = 20)
```



```
## Using `ggtitle()`, `xlab()`, and `ylab()` to add a title, x label, and y label to the previous plot
## Title: HSGegree by population
## X label: HSGegree percentage
## Y Label: county count
title <- 'HSGegree by population'
x_label <- 'HSGegree percentage'
y_label <- 'county count'
ggplot(us_census_df, aes(HSDegree)) + geom_histogram( bins = 20) +
  ggtitle(title) + labs(x_label, y_label)
```

HSGegree by population

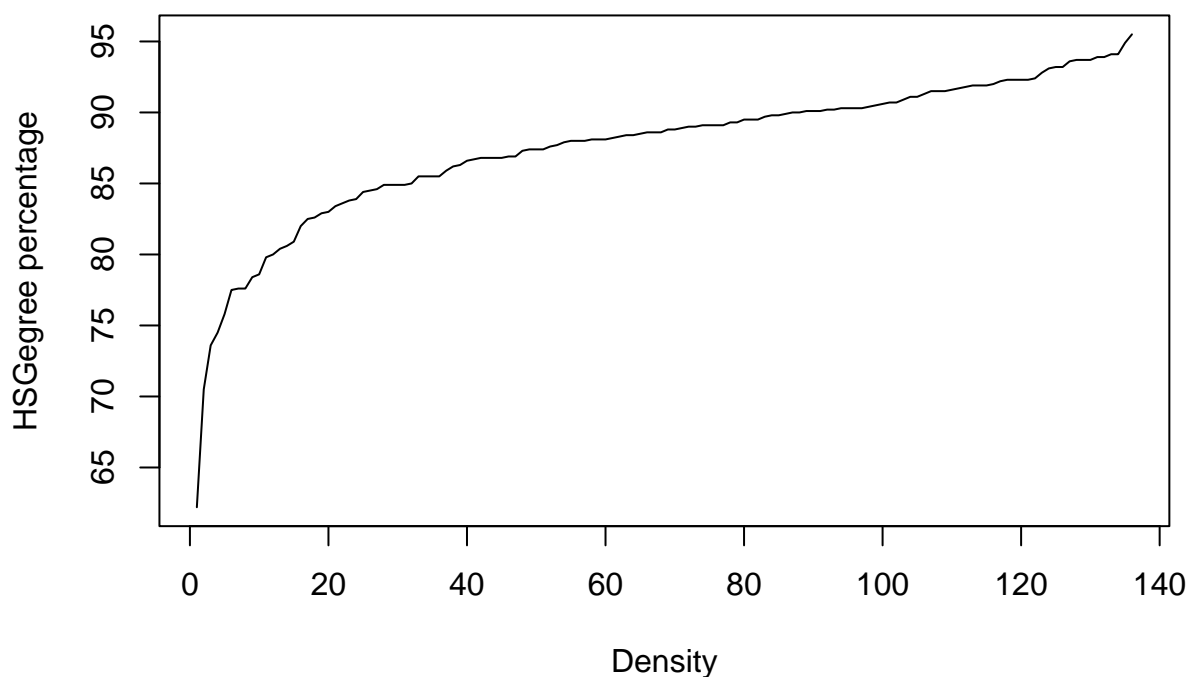


```
## Probability plot
HSDegree <- us_census_df$HSDegree
HSDegree_sort <- sort(HSDegree, decreasing = FALSE)

plot(x = HSDegree_sort, type= "l",
     xlab = "Density",
     ylab = "HSGegree percentage",
     main = "HSGegree by population"
)
```



## HSegree by population



```
## stat.desc()
library (pastecs)
```

```
## Warning: package 'pastecs' was built under R version 4.2.3
```

```
stat.desc(us_census_df$HSDegree, basic=TRUE, desc=TRUE, norm=TRUE)
```

##	nbr.val	nbr.null	nbr.na	min	max
##	1.360000e+02	0.000000e+00	0.000000e+00	6.220000e+01	9.550000e+01
##	range	sum	median	mean	SE.mean
##	3.330000e+01	1.191800e+04	8.870000e+01	8.763235e+01	4.388598e-01
##	CI.mean.0.95	var	std.dev	coef.var	skewness
##	8.679296e-01	2.619332e+01	5.117941e+00	5.840241e-02	-1.674767e+00
##	skew.2SE	kurtosis	kurt.2SE	normtest.W	normtest.p
##	-4.030254e+00	4.352856e+00	5.273885e+00	8.773635e-01	3.193634e-09