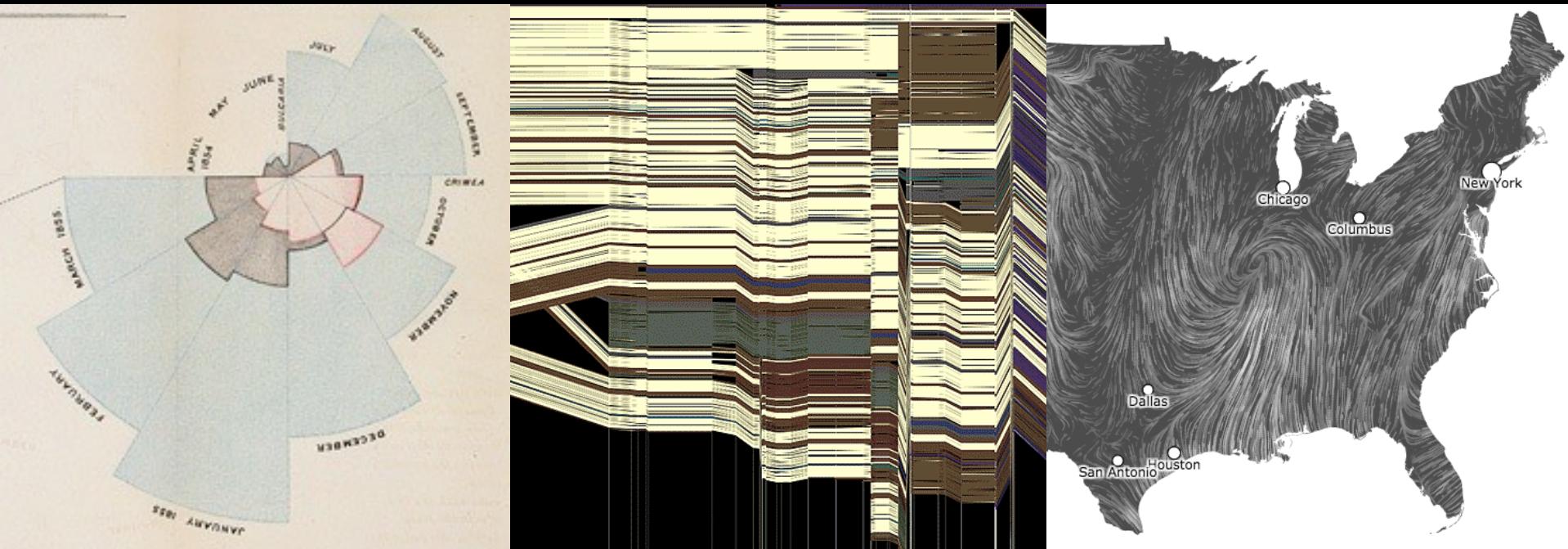


CSE 512 - Data Visualization

Design & Re-Design



Jeffrey Heer University of Washington

Last Time:
Data & Image Models

The Big Picture

task
questions, goals
assumptions

data
physical data type
conceptual data type

domain
metadata
semantics
conventions

processing
algorithms

mapping
visual encoding

image
visual channel
graphical marks

Nominal, Ordinal & Quantitative

N - Nominal (labels or categories)

- Operations: $=, \neq$

O - Ordered

- Operations: $=, \neq, <, >$

Q - Interval (location of zero arbitrary)

- Operations: $=, \neq, <, >, -$
- Can measure distances or spans

Q - Ratio (zero fixed)

- Operations: $=, \neq, <, >, -, \%$
- Can measure ratios or proportions

Visual Encoding Variables

Position (x 2)
Size
Value
Texture
Color
Orientation
Shape

Others?

		LES VARIABLES DE L'IMAGE			
		POINTS	LIGNES	ZONES	
Z	XY 2 DIMENSIONS DU PLAN	x	x	x	
	TAILLE			-	
	VALEUR				
		LES VARIABLES DE SÉPARATION DES IMAGES			
		GRAIN			
		COULEUR			
		ORIENTATION			
		FORME			

Bertin's “Levels of Organization”

Position

N	O	Q
---	---	---

Nominal

Size

N	O	Q
---	---	---

Ordinal

Value

N	O	Q
---	---	---

Quantitative

Texture

N	o	
---	---	--

Color

N		
---	--	--

Orientation

N		
---	--	--

Shape

N		
---	--	--

Note: **Q** ⊂ **O** ⊂ **N**

Choosing Visual Encodings

Assume k visual encodings and n data attributes.
We would like to pick the “best” encoding among
a combinatorial set of possibilities of size $(n+1)^k$

Principle of Consistency

The properties of the image (visual variables)
should match the properties of the data.

Principle of Importance Ordering

Encode the most important information in the
most effective way.

Design Criteria [Mackinlay 86]

Expressiveness

A set of facts is *expressible* in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

Effectiveness

A visualization is more *effective* than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.

Design Criteria *Translated*

Tell the truth and nothing but the truth
(don't lie, and don't lie by omission)

Use encodings that people decode better
(where better = faster and/or more accurate)

Effectiveness Rankings

[Mackinlay 86]

QUANTITATIVE

Position
Length
Angle
Slope
Area (Size)
Volume
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Shape

ORDINAL

Position
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Length
Angle
Slope
Area (Size)
Volume
Shape

NOMINAL

Position
Color Hue
Texture
Connection
Containment
Density (Value)
Color Sat
Shape
Length
Angle
Slope
Area
Volume

Effectiveness Rankings

[Mackinlay 86]

QUANTITATIVE

Position

Length

Angle

Slope

Area (Size)

Volume

Density (Value)

Color Sat

Color Hue

Texture

Connection

Containment

Shape

ORDINAL

Position

Density (Value)

Color Sat

Color Hue

Texture

Connection

Containment

Length

Angle

Slope

Area (Size)

Volume

Shape

NOMINAL

Position

Color Hue

Texture

Connection

Containment

Density (Value)

Color Sat

Shape

Length

Angle

Slope

Area

Volume

Effectiveness Rankings

[Mackinlay 86]

QUANTITATIVE

Position
Length
Angle
Slope
Area (Size)
Volume
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Shape

ORDINAL

Position
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Length
Angle
Slope
Area (Size)
Volume
Shape

NOMINAL

Position
Color Hue
Texture
Connection
Containment
Density (Value)
Color Sat
Shape
Length
Angle
Slope
Area
Volume

A1 Review

A1 Submission Designs

Measures: Population count, Age distribution, Gender ratio, Gender difference, Growth rate

Transforms: Percentages, Counts, Proportions

Encodings: Bar, Line/Area, Pie, Dot/Scatter, Hybrid

Bars: Stacked, Grouped, Opposed Axes

Extra Context: More Years, Mortality, Race, ...

Design Considerations

Title, labels, legend, captions, source!

Expressiveness and Effectiveness

Avoid unexpressive marks (lines? gradients?)

Use perceptually effective encodings

Don't distract: faint gridlines, pastel highlights/fills

The "elimination diet" approach – start minimal

Support comparison and pattern perception

Between elements, to a reference line, or to totals

Design Considerations

Transform data (e.g., invert, log, normalize)

Group / sort data by meaningful dimensions

Reduce cognitive overhead

Minimize visual search, minimize ambiguity

Appropriate size, aspect ratio, legible text

Avoid legend lookups if direct labeling works

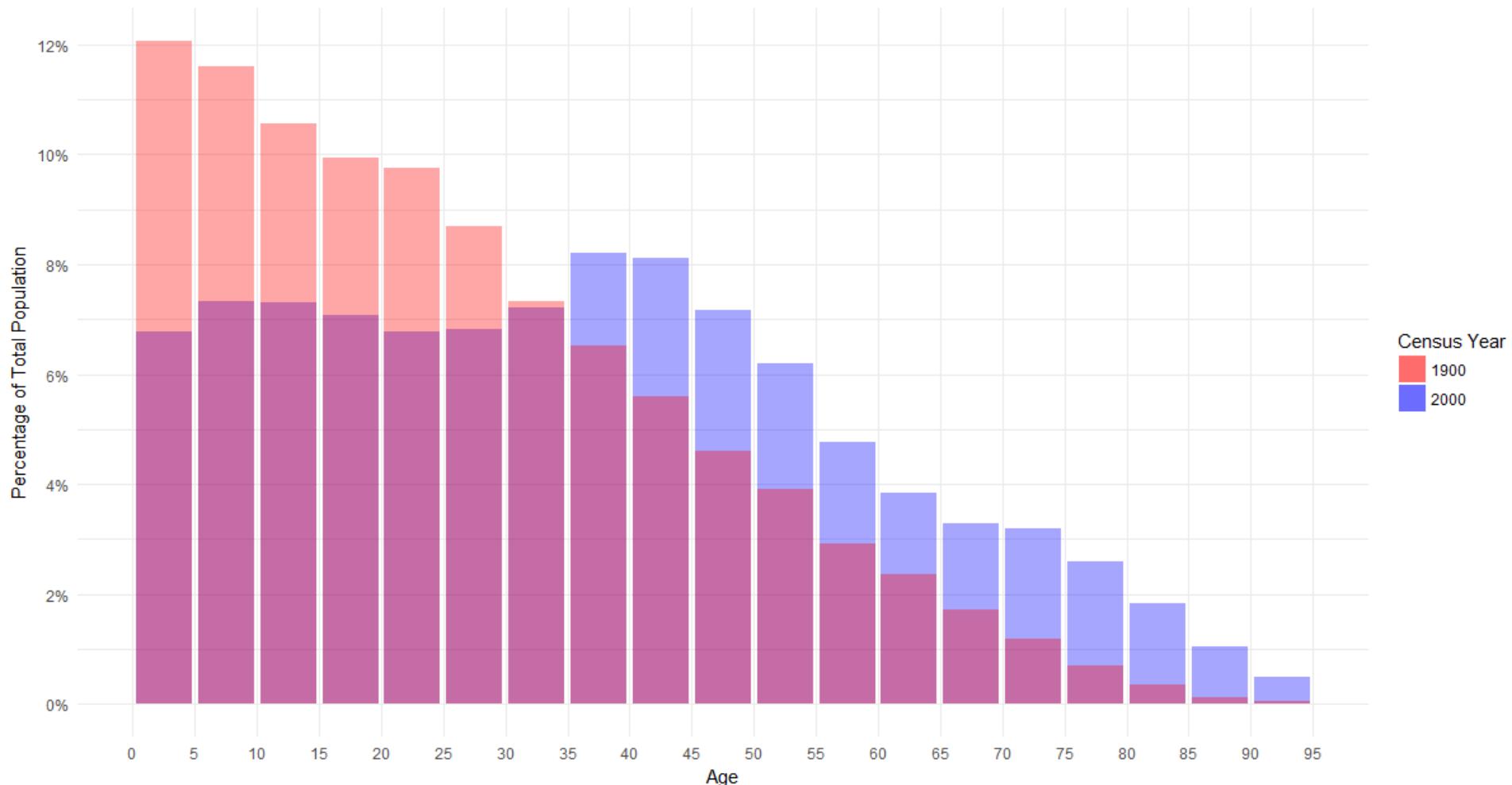
Avoid color mappings with indiscernible colors

Be consistent! Visual inferences should consistently support data inferences.

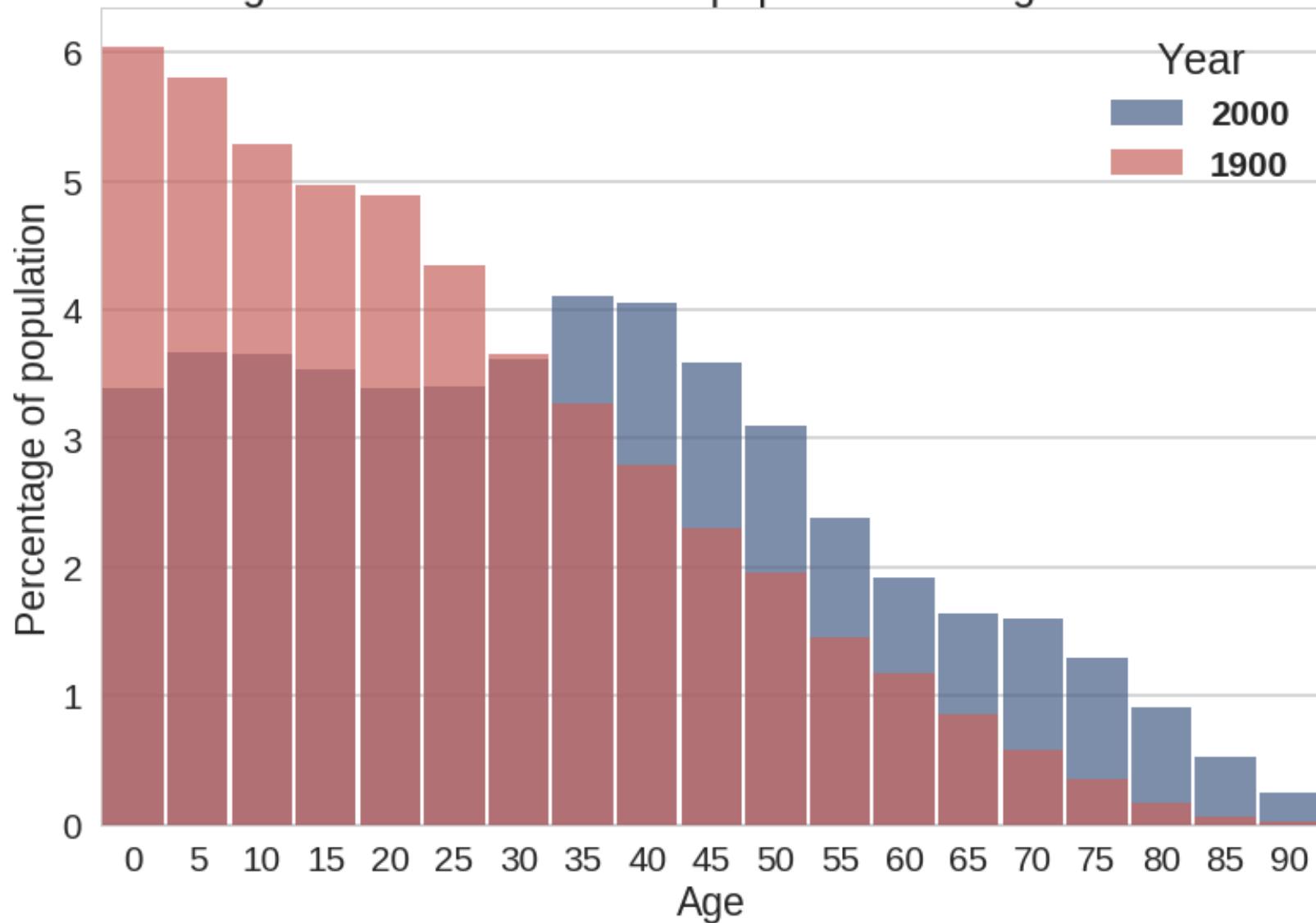
A1 Submissions

Bar Charts

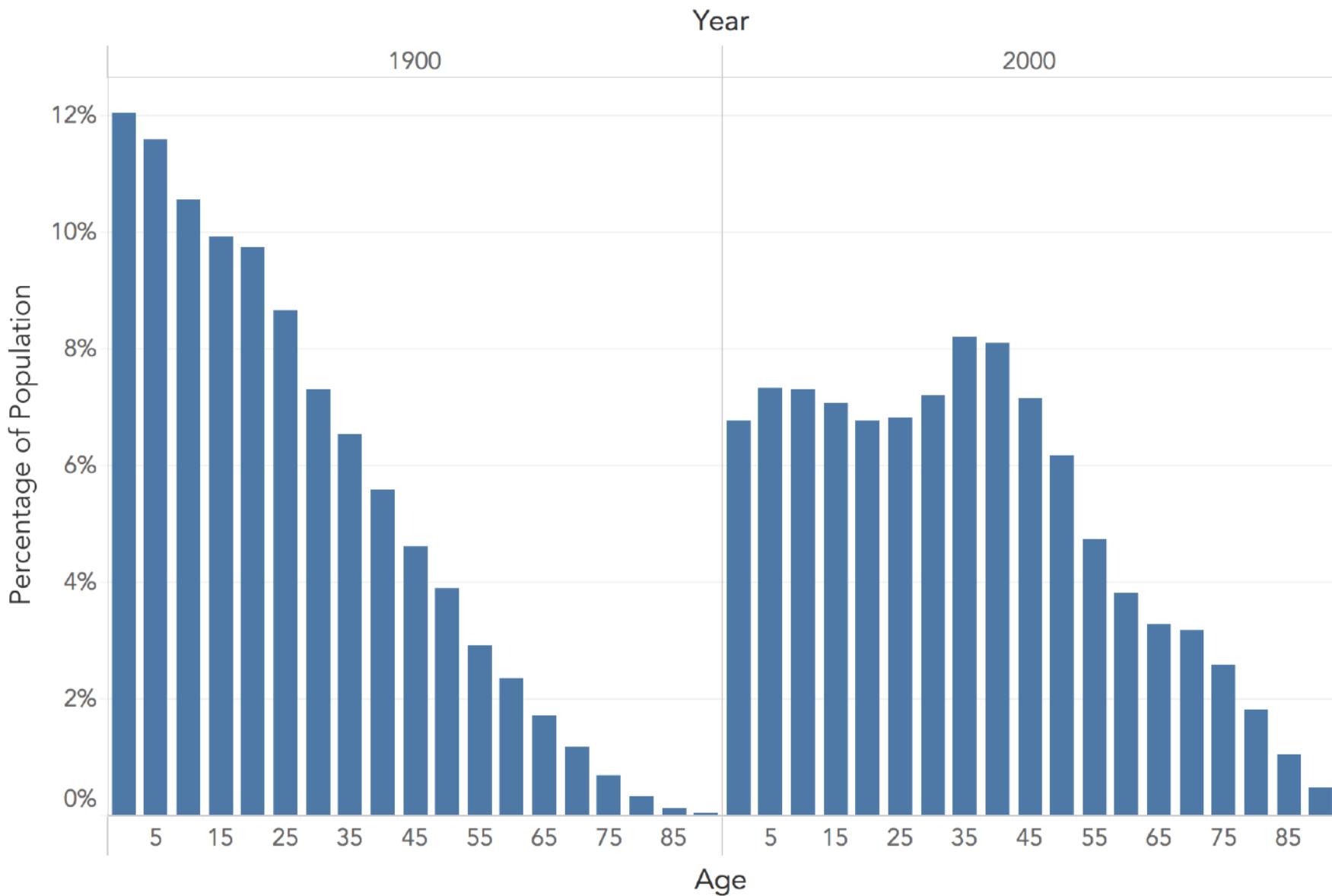
Was the Population of the United States Older in 2000 Than It Was in 1900?



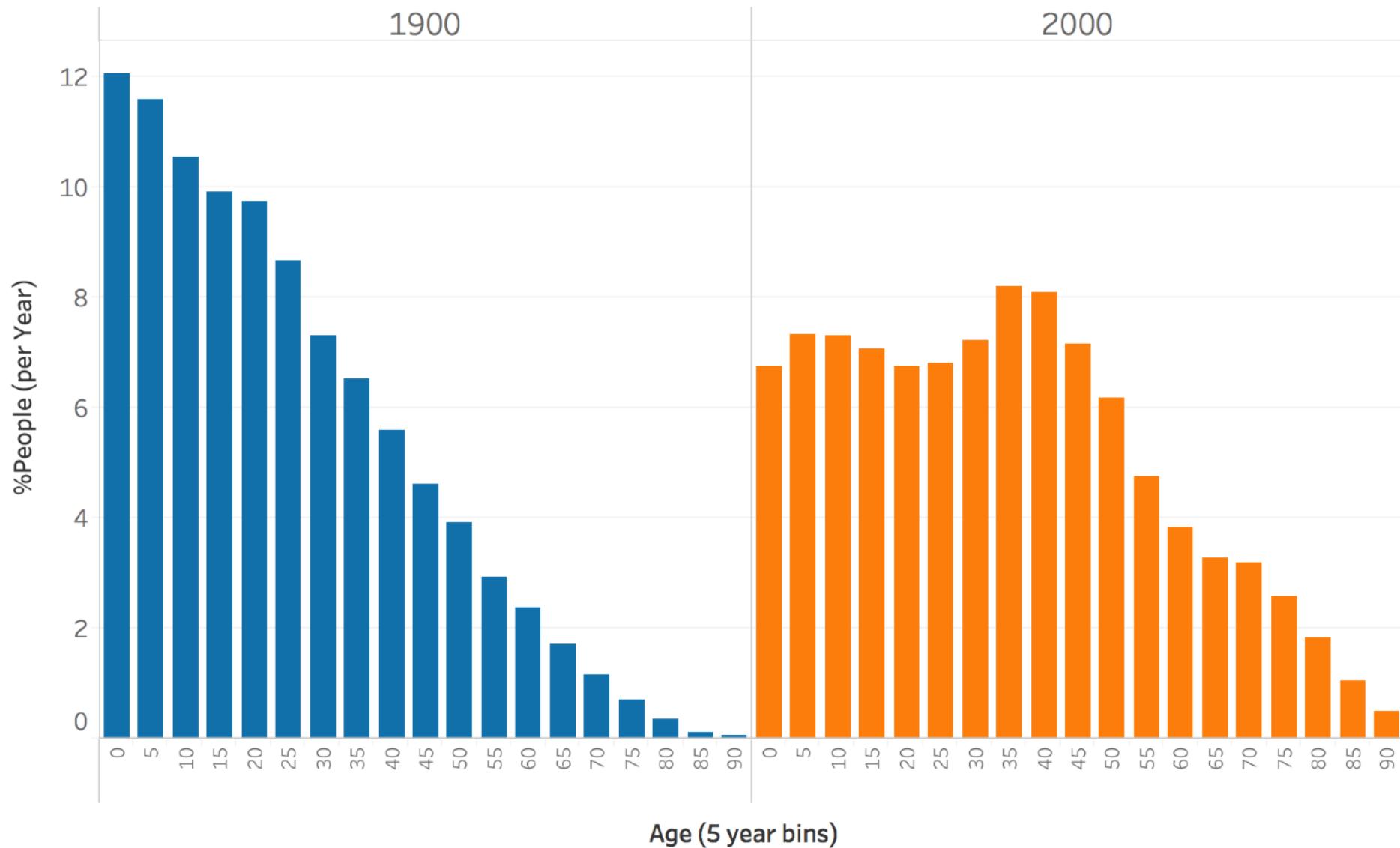
How did the age distribution of the US population change from 1900 to 2000?



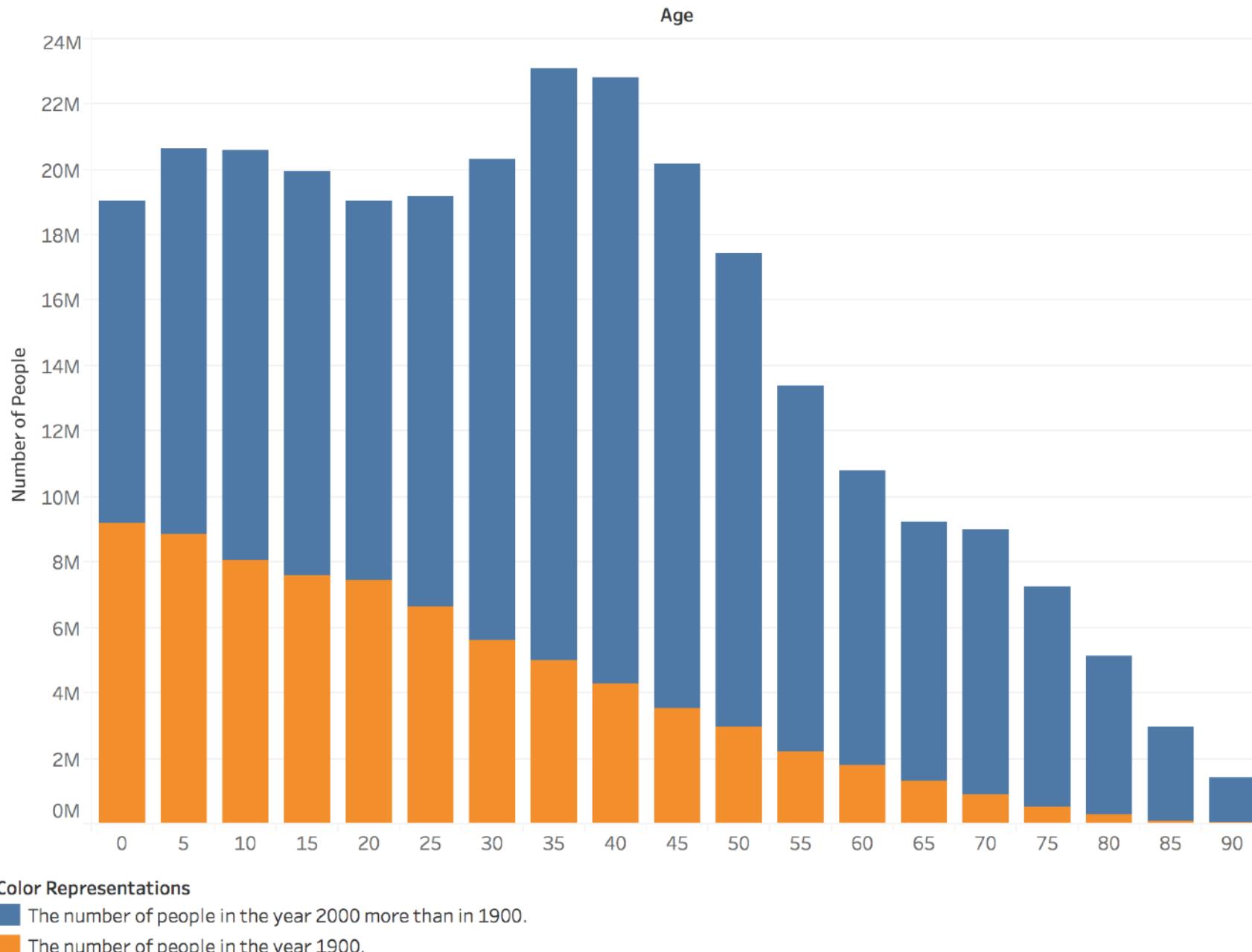
How has the distribution of age changed in the United States over the past century?



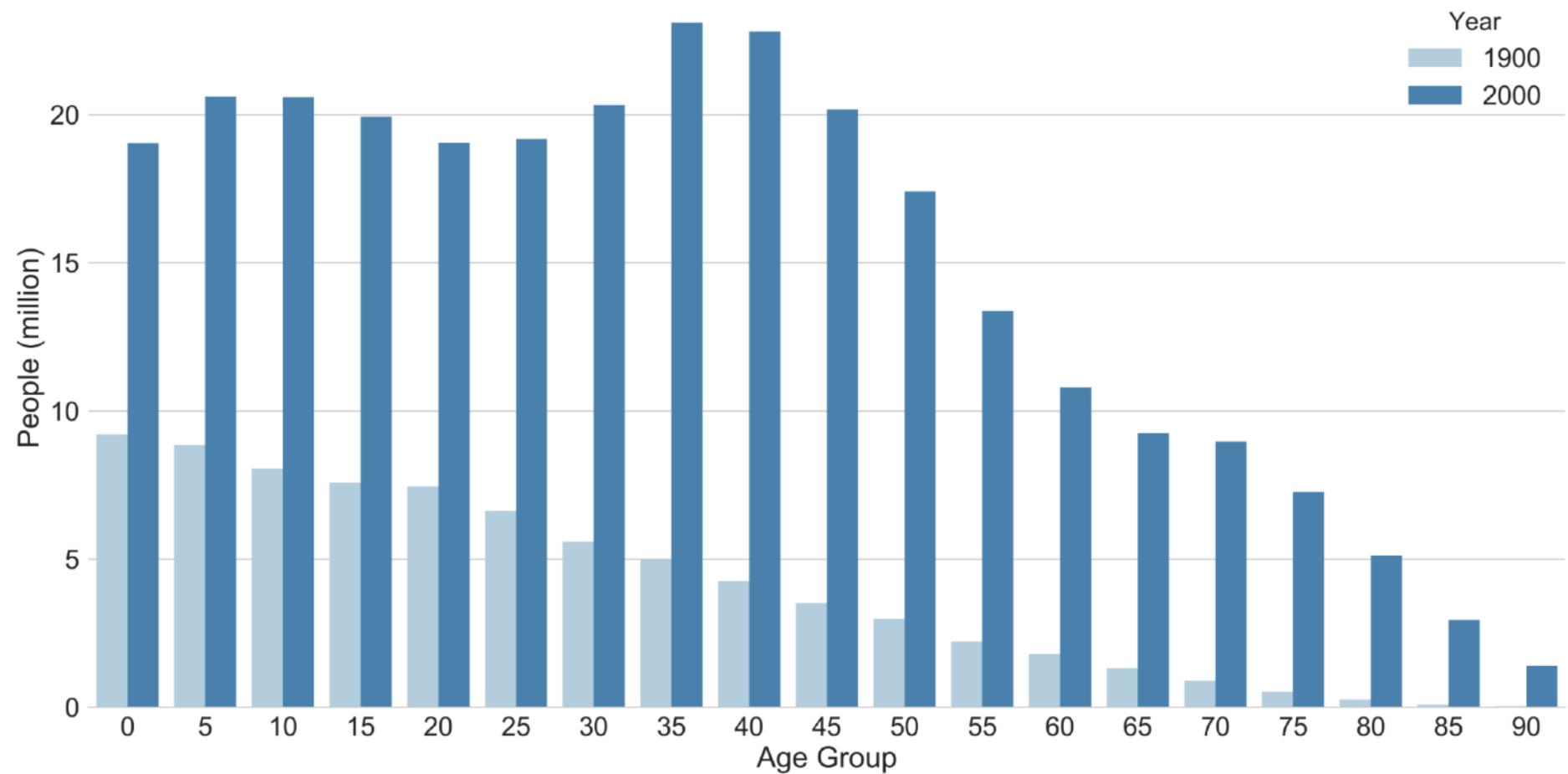
How have population age distributions changed between 1900 and 2000?



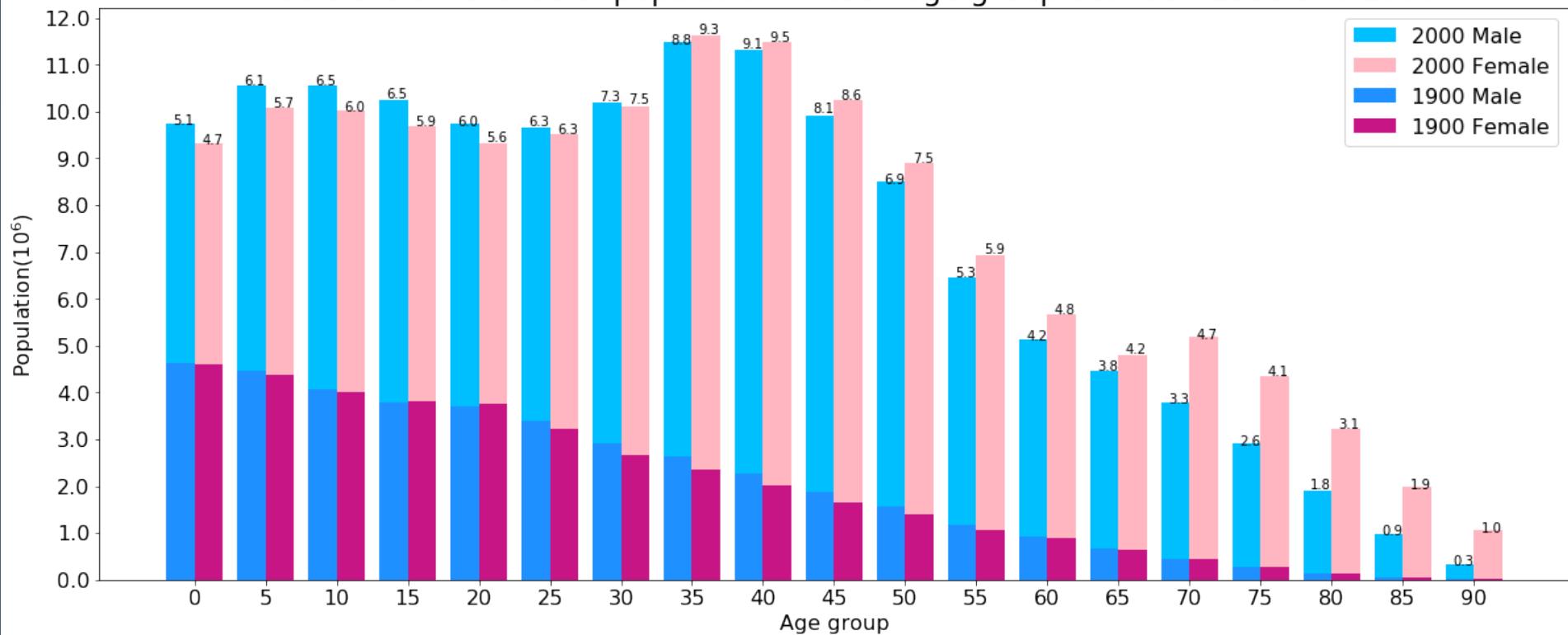
What is the difference in the number of people in the year 1900 and 2000 at different age groups?



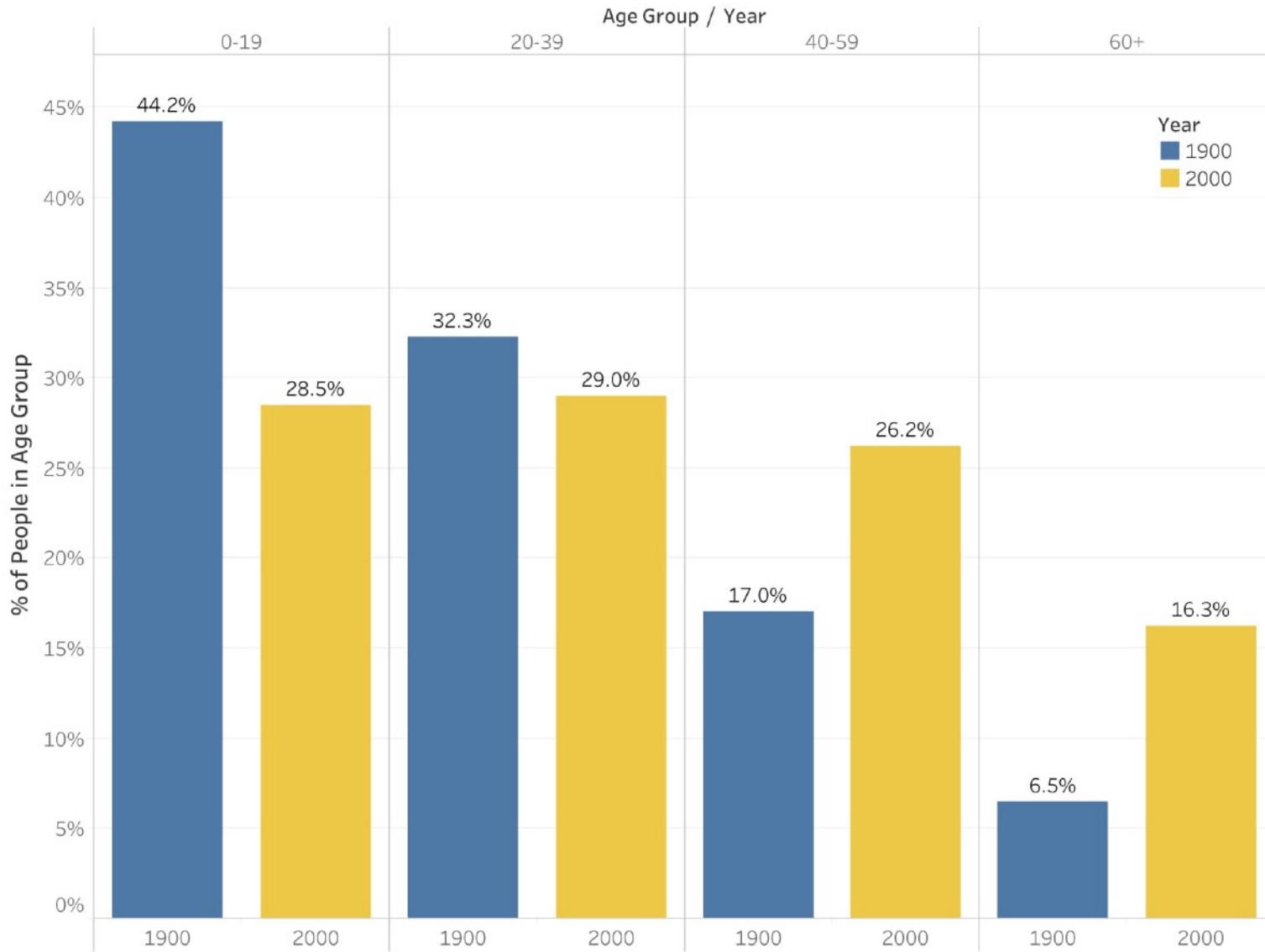
What is the change in population distribution by age from 1900 to 2000?



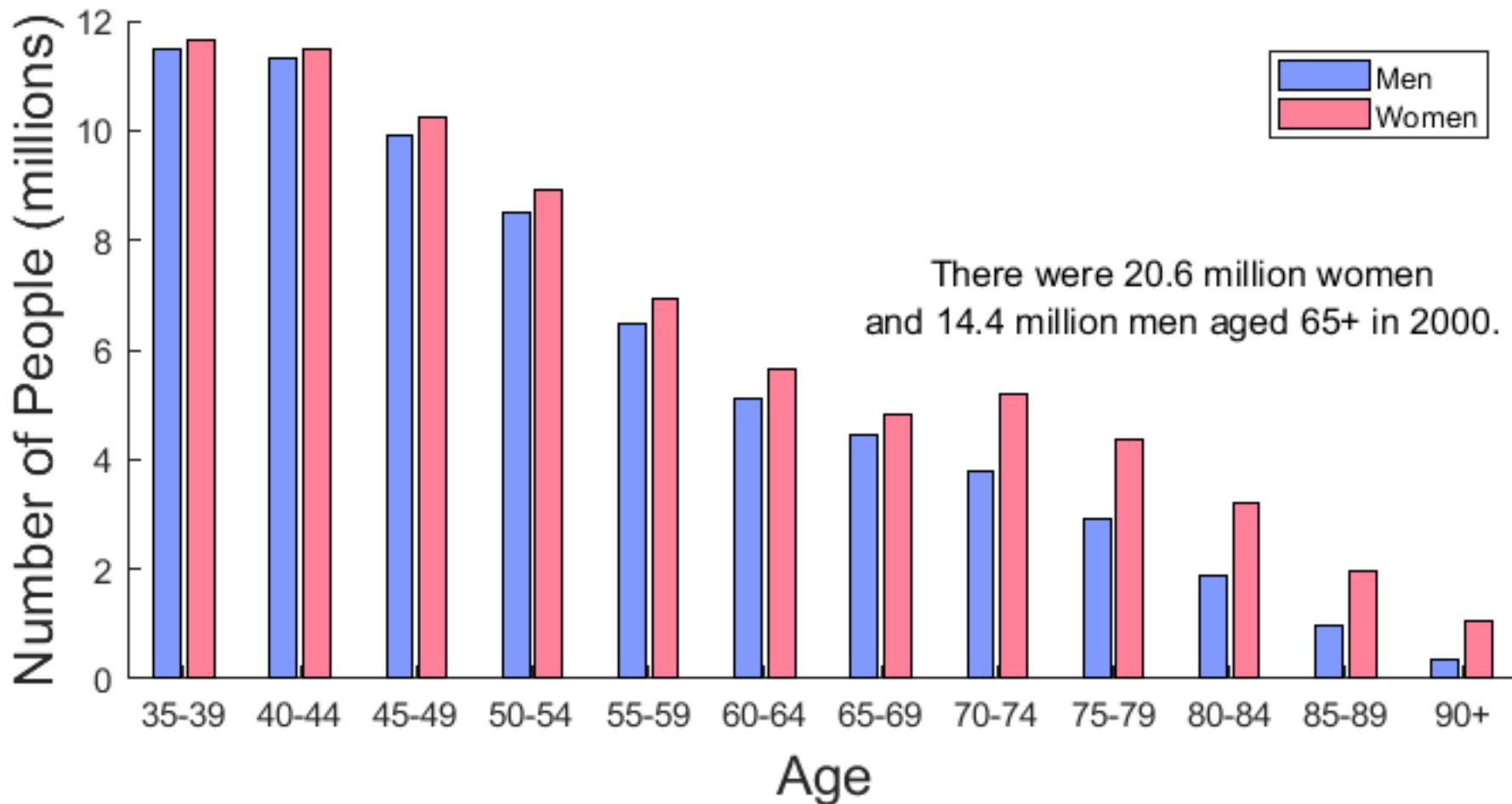
What is the difference of population in each age group between 1900 and 2000



How is the US population distributed by age groups in 1900 vs. 2000?

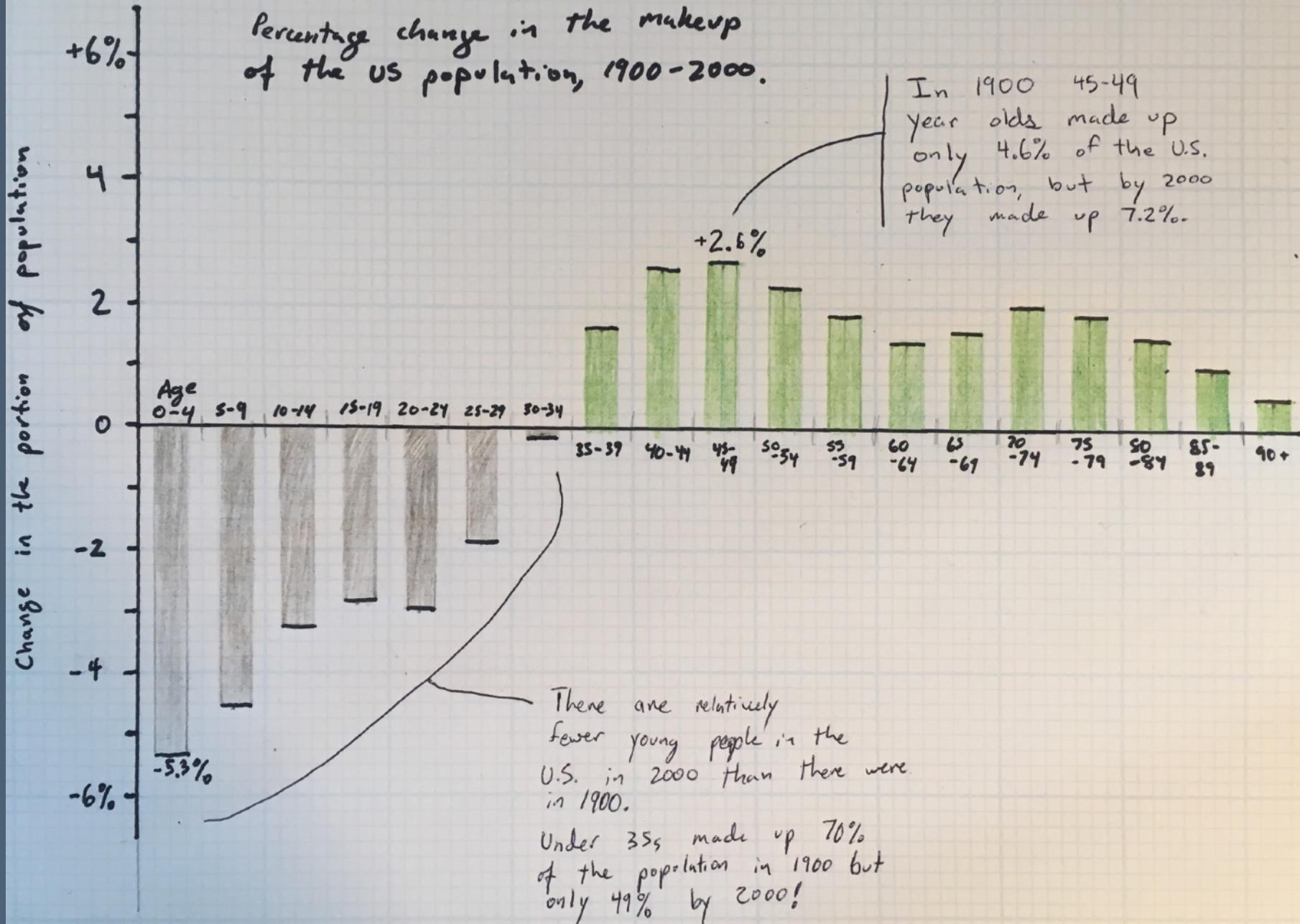


Who made it to retirement in 2000?



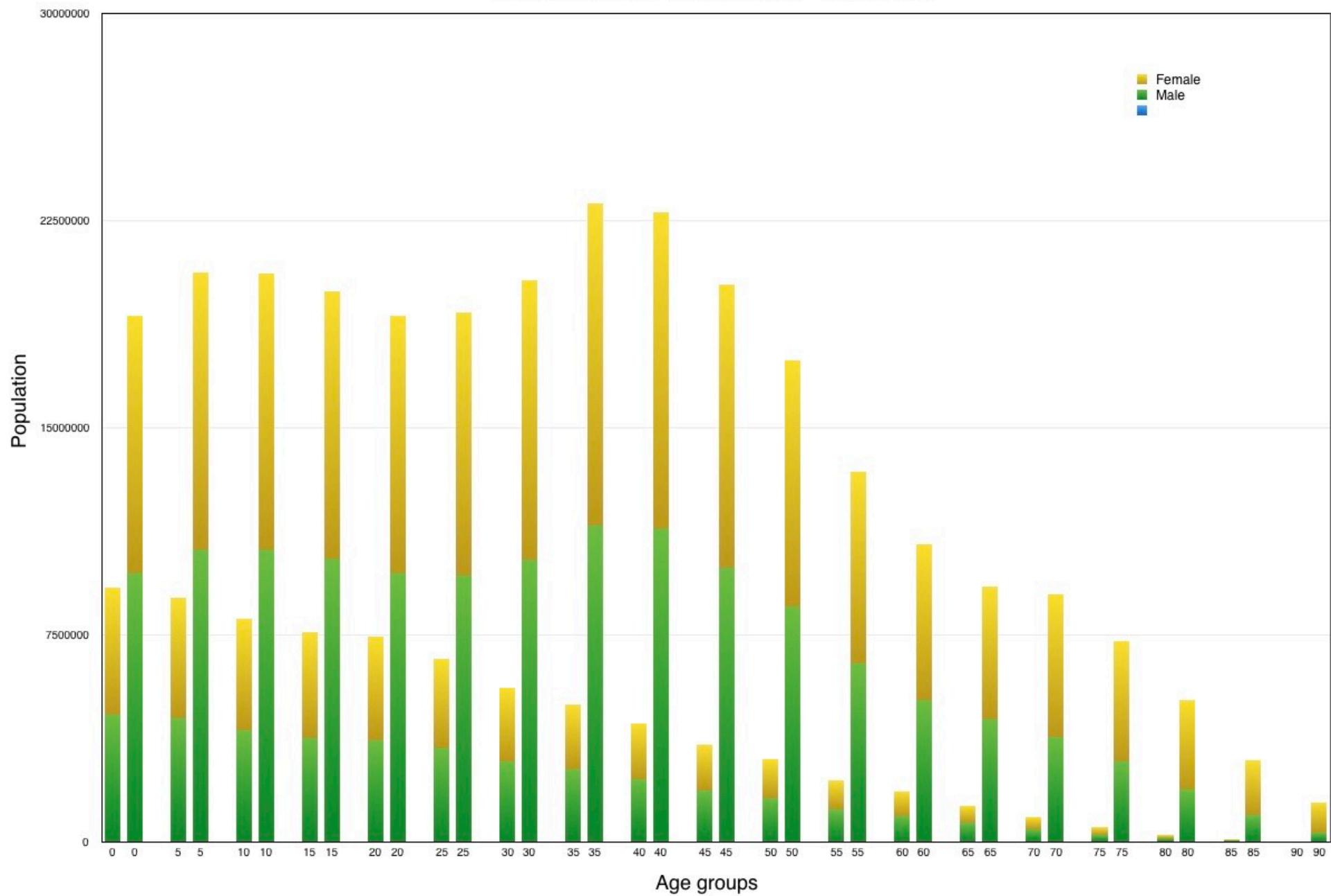
The Aging Population

Percentage change in the makeup
of the US population, 1900-2000.

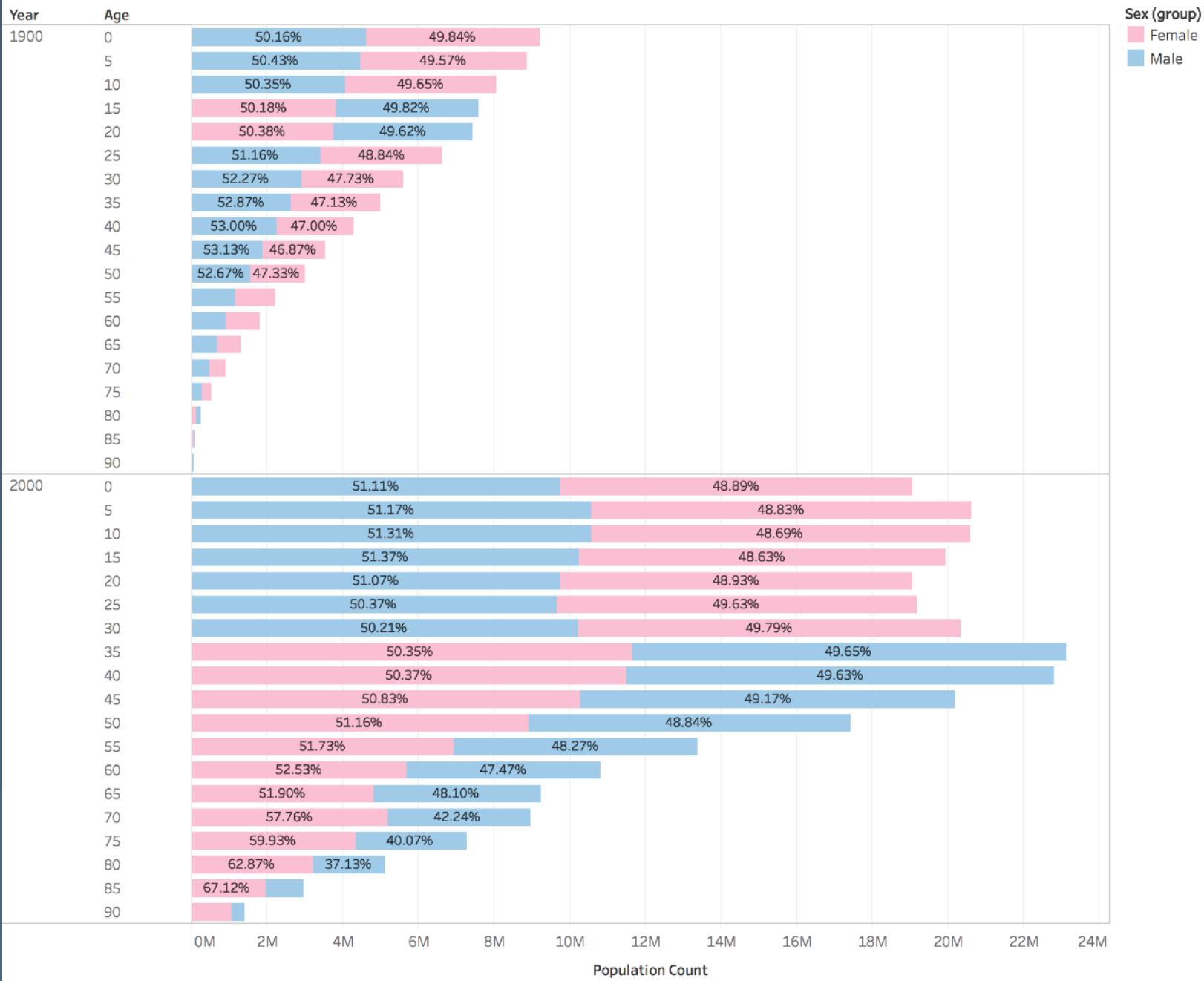


Stacked Bars

Comparative study of census data - 1900 vs 2000

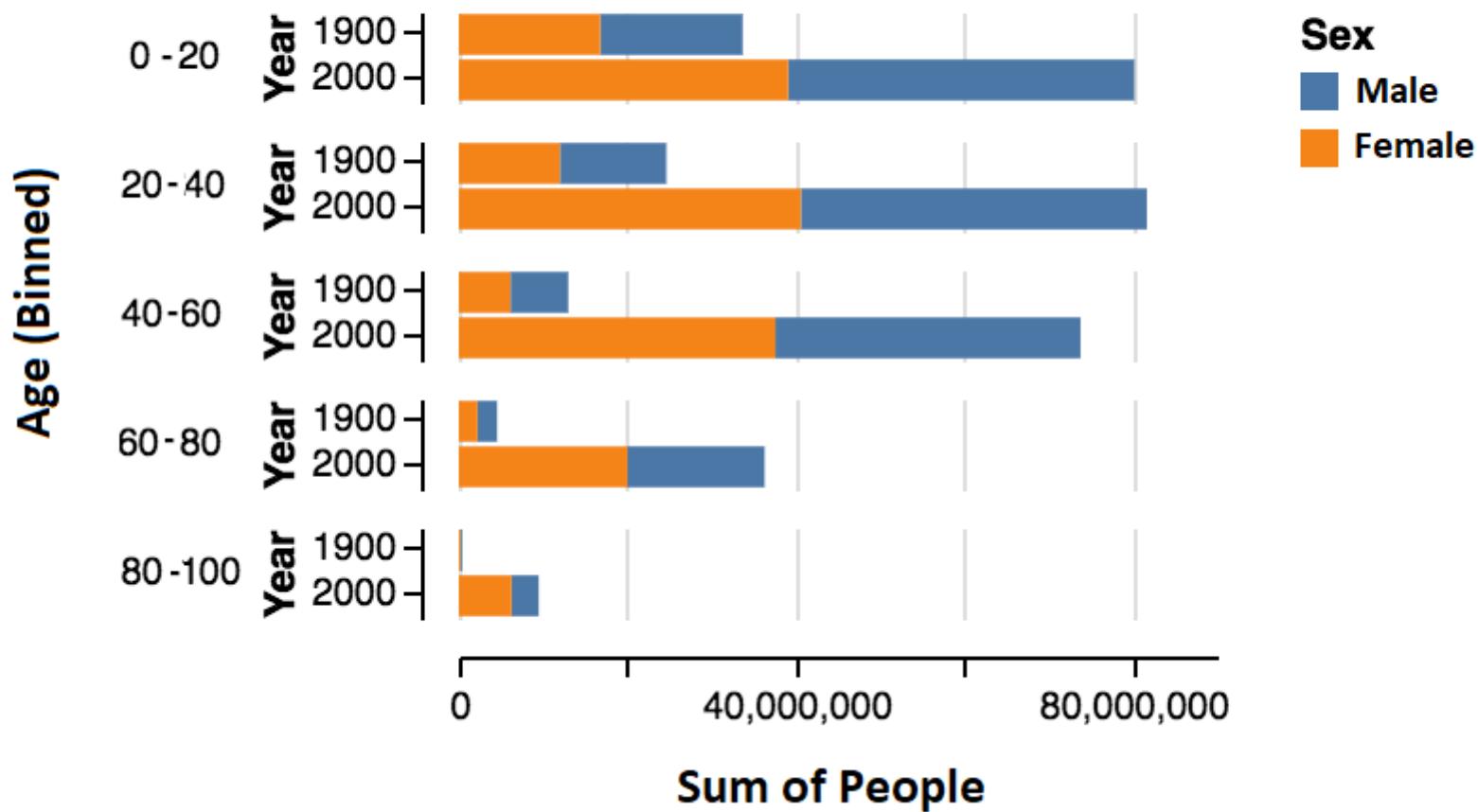


U.S. Population By 5-Year Age Group & Gender: Census 1900 vs 2000

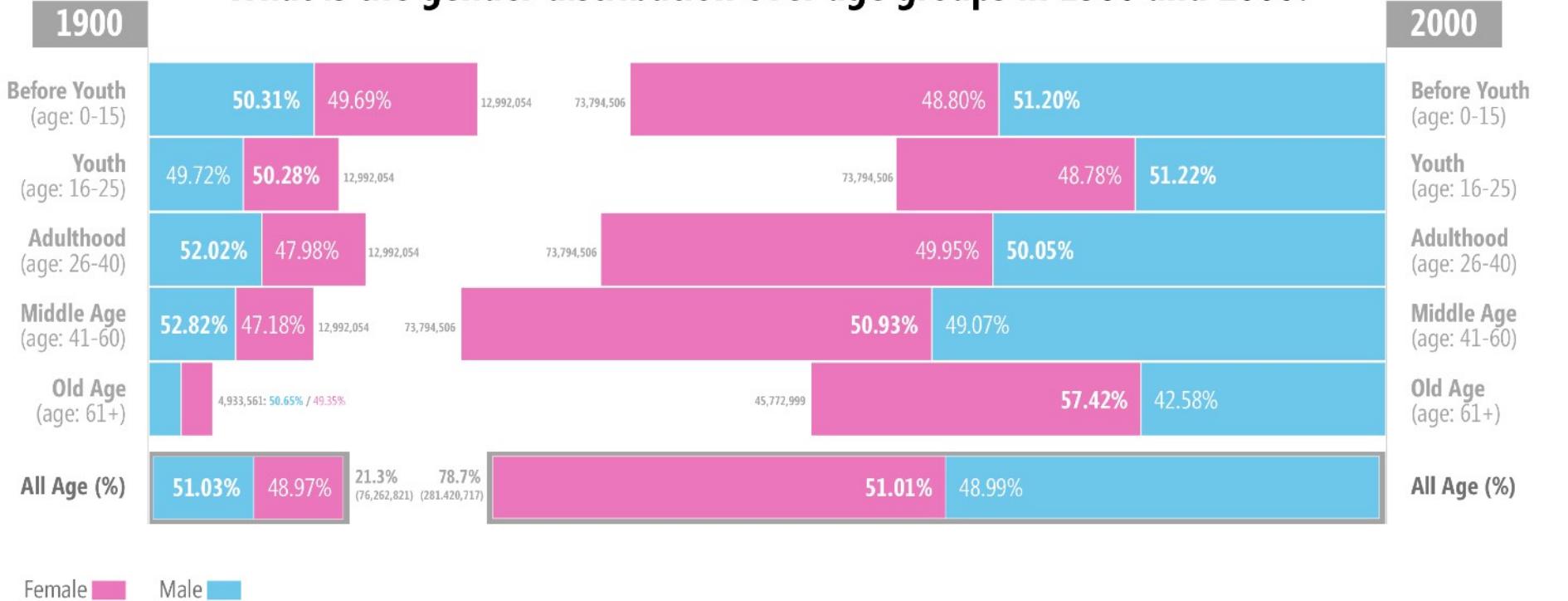


Sum of People for each Age broken down by Year. Color shows details about Sex (group). The marks are labeled by Prop. Details are shown for Sex (group).

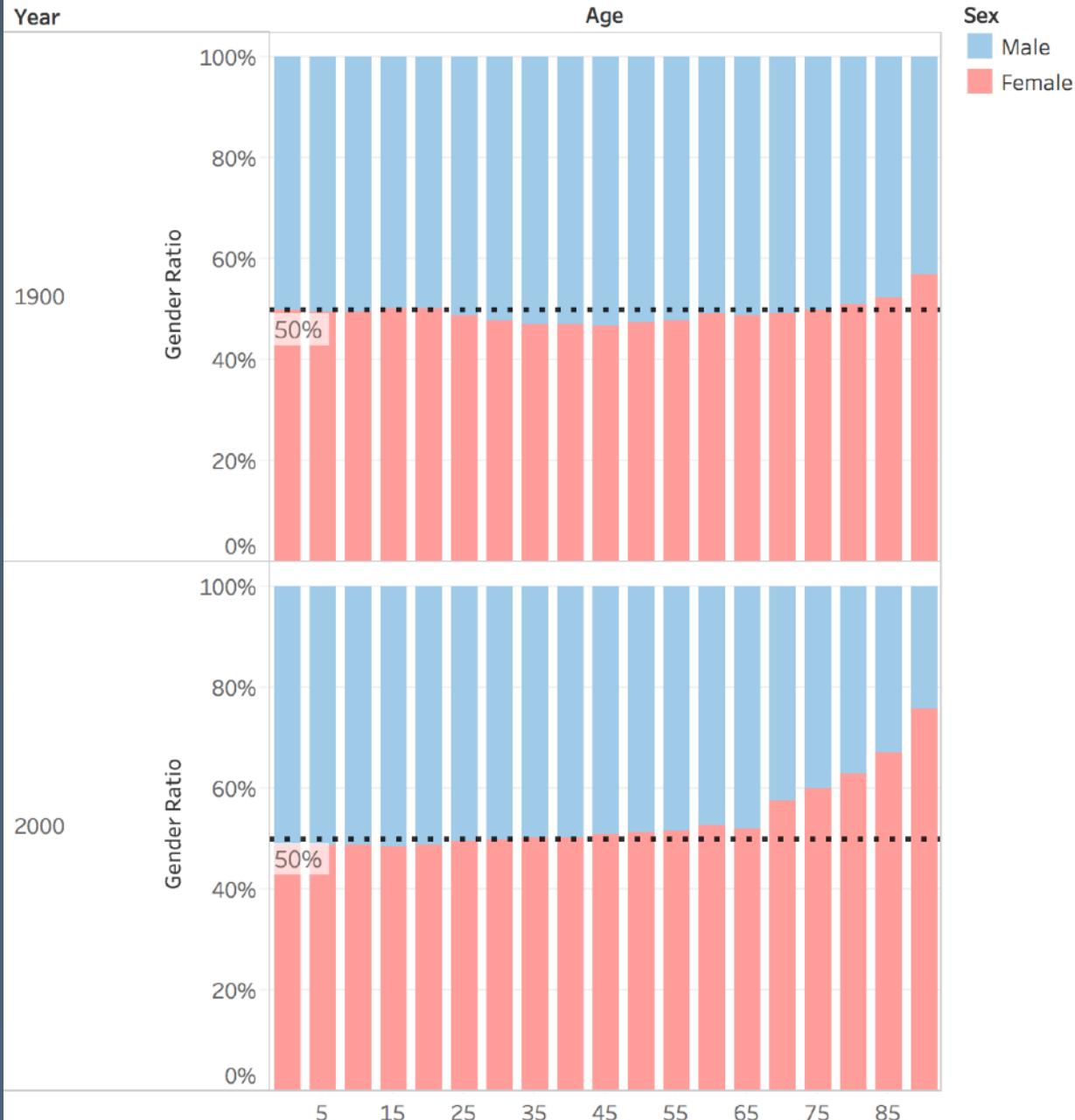
How have the demographics of male and females changed from 1900 to 2000?



What is the gender distribution over age groups in 1900 and 2000?

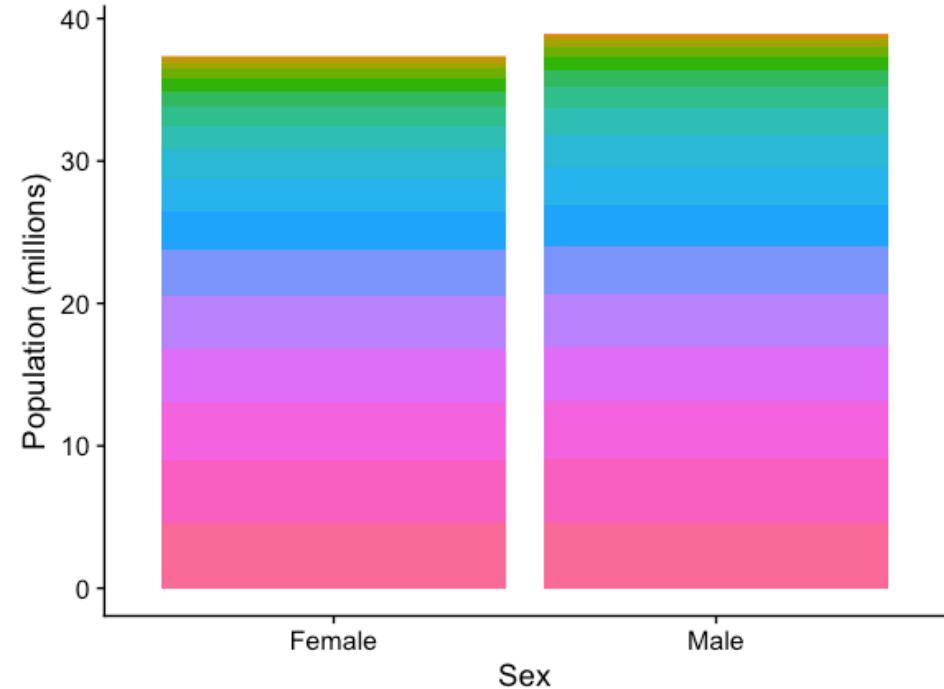


Is gender ratio balanced across different years and ages?

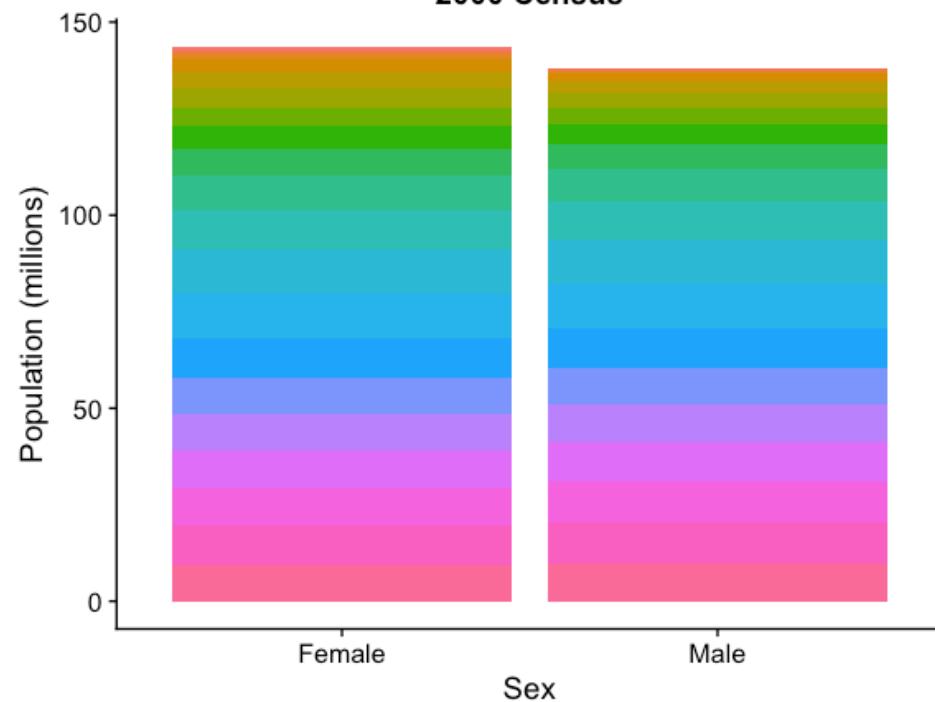


Comparing Age and Sex Demographics in the United States in 1900 and 2000

1900 Census



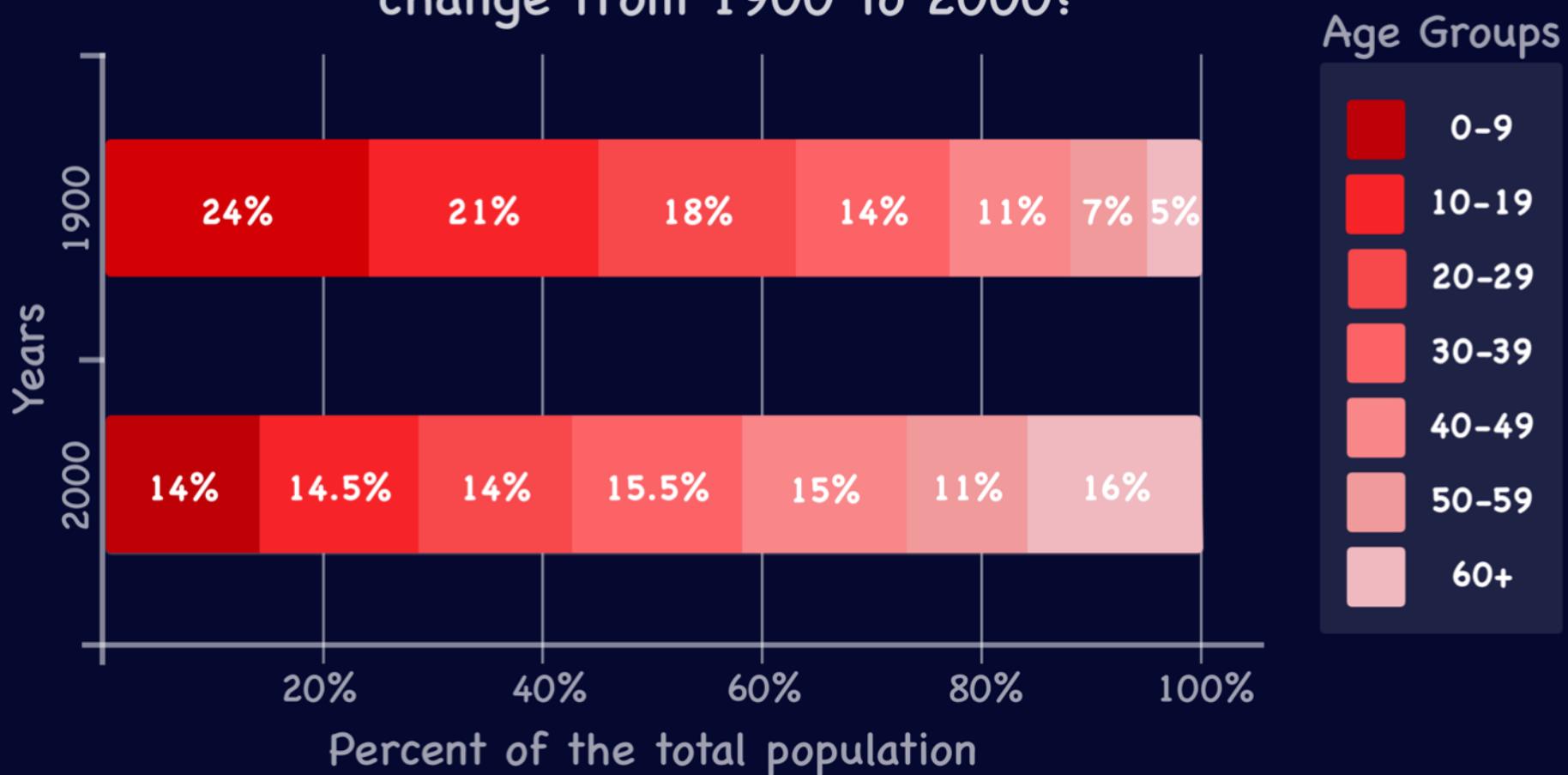
2000 Census



Age Groups

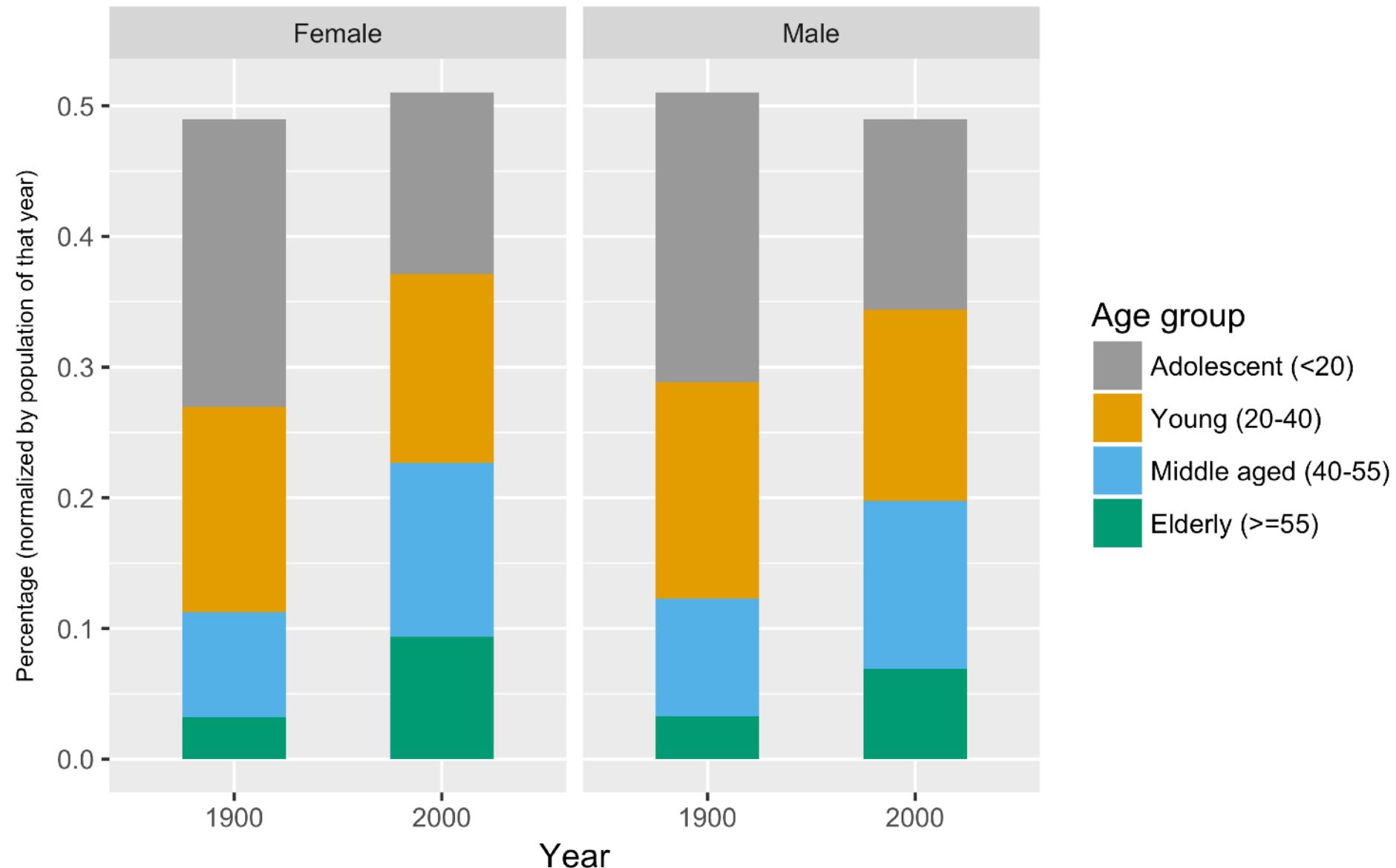
0 - 4	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74	75 - 79	80 - 84	85 - 89	90 +
pink	light pink	purple	light purple	blue	light blue	teal	light teal	green	light green	yellow-green	yellow	orange-yellow	orange	light orange	yellow-orange	orange-red	red	dark red

How does the age distribution of U.S. population change from 1900 to 2000?

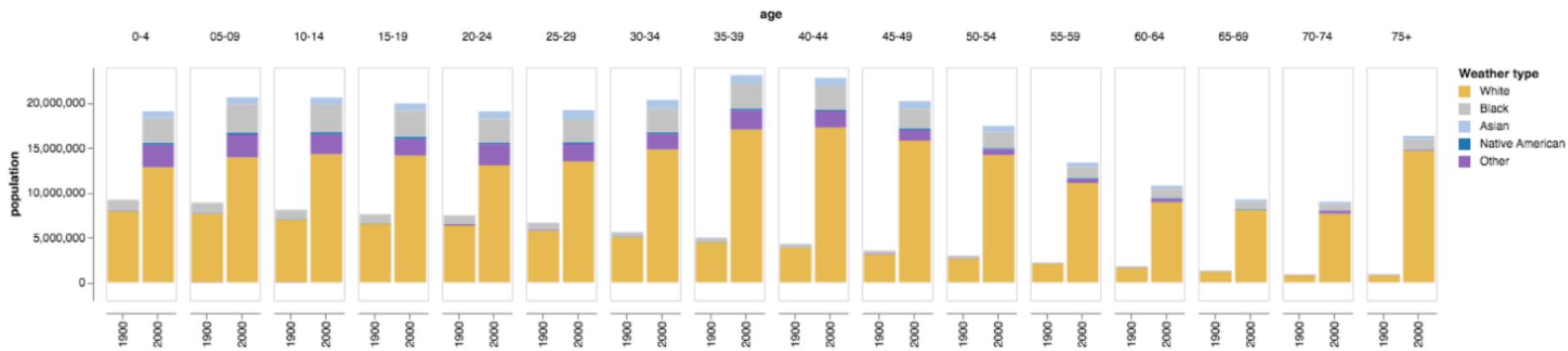


Age composition change of U.S. Population over time by sex

via census data

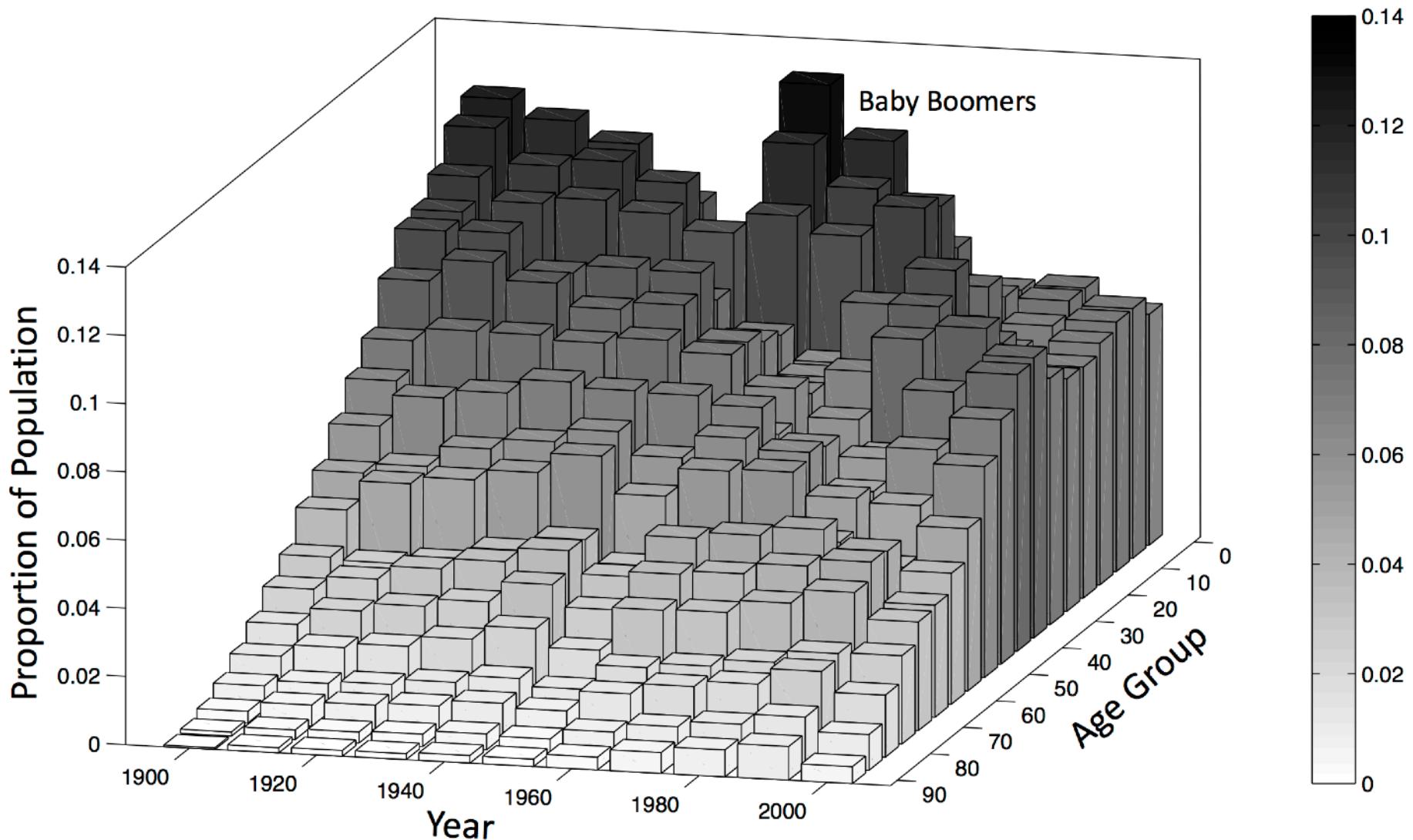


How does the racial composition of the US change between 1900 and 2000?

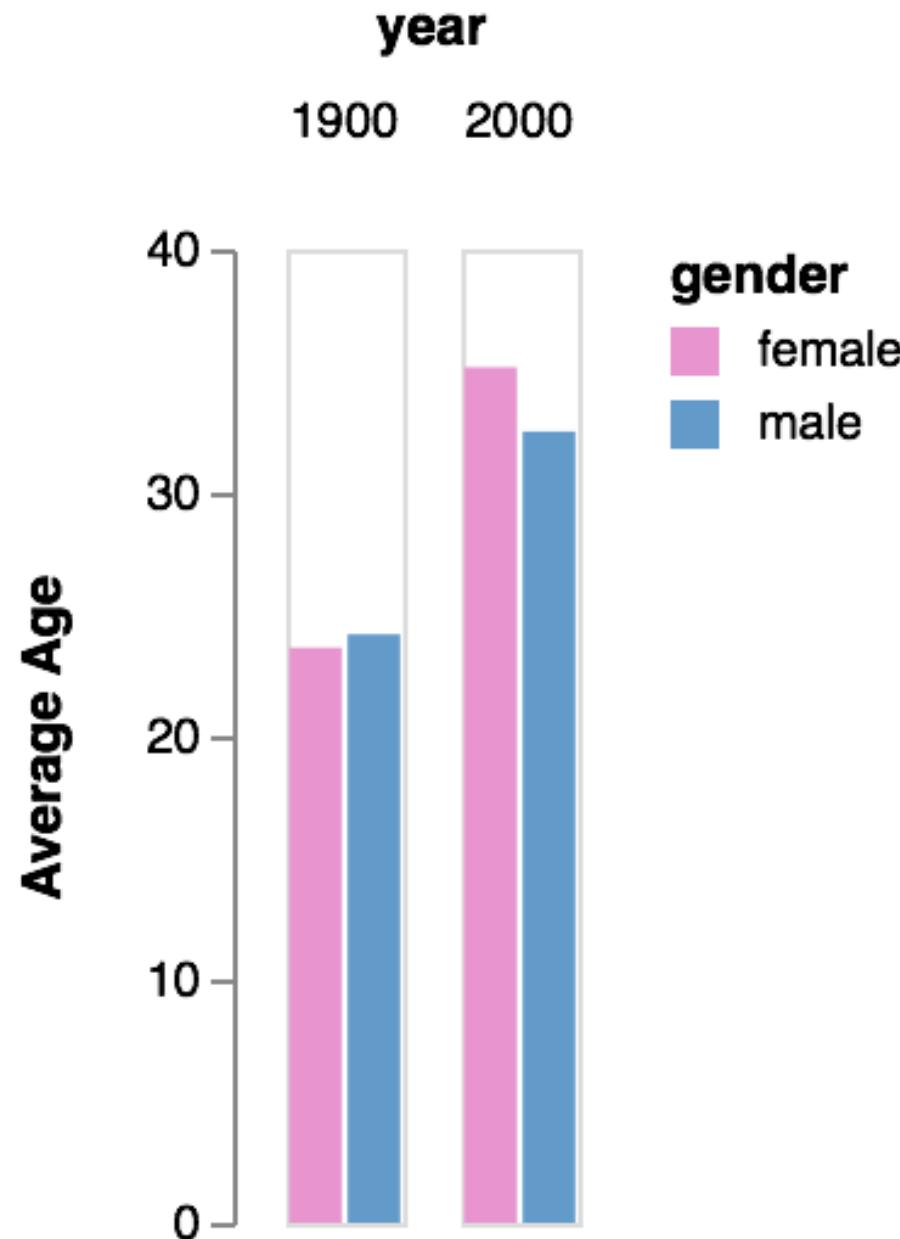


The 3rd Dimension

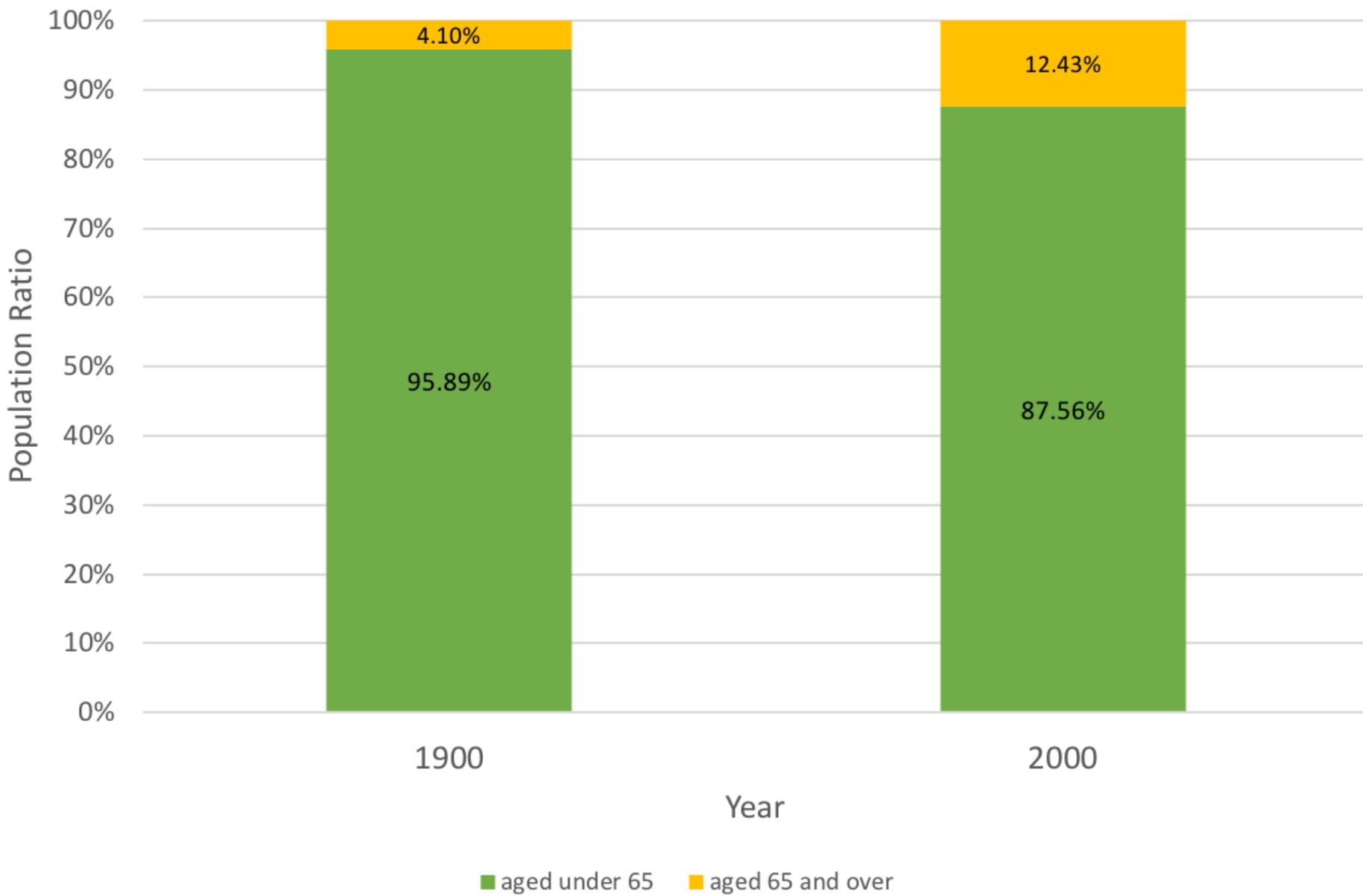
How Did the U.S. Age Distribution Change from 1900 to 2000?



Simplify!



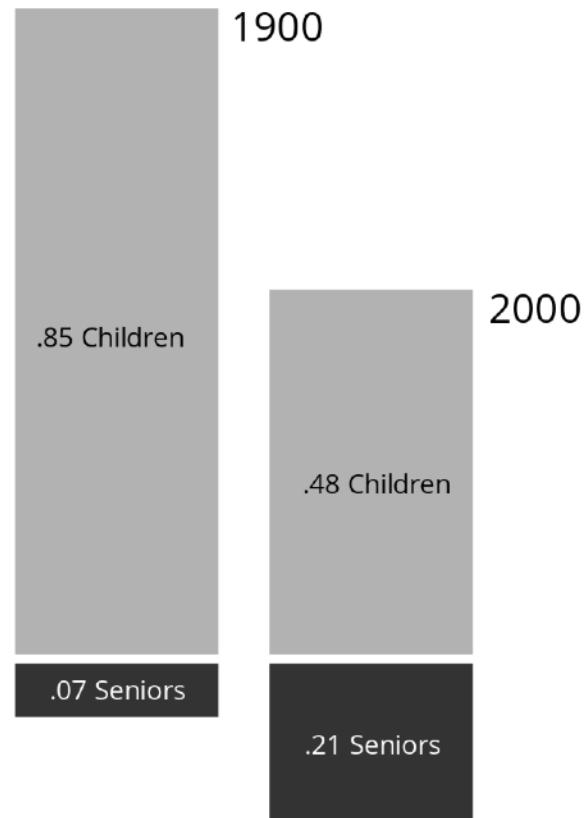
Population Ageing



How have populations of males and females changed with respect to each other over 100 years?

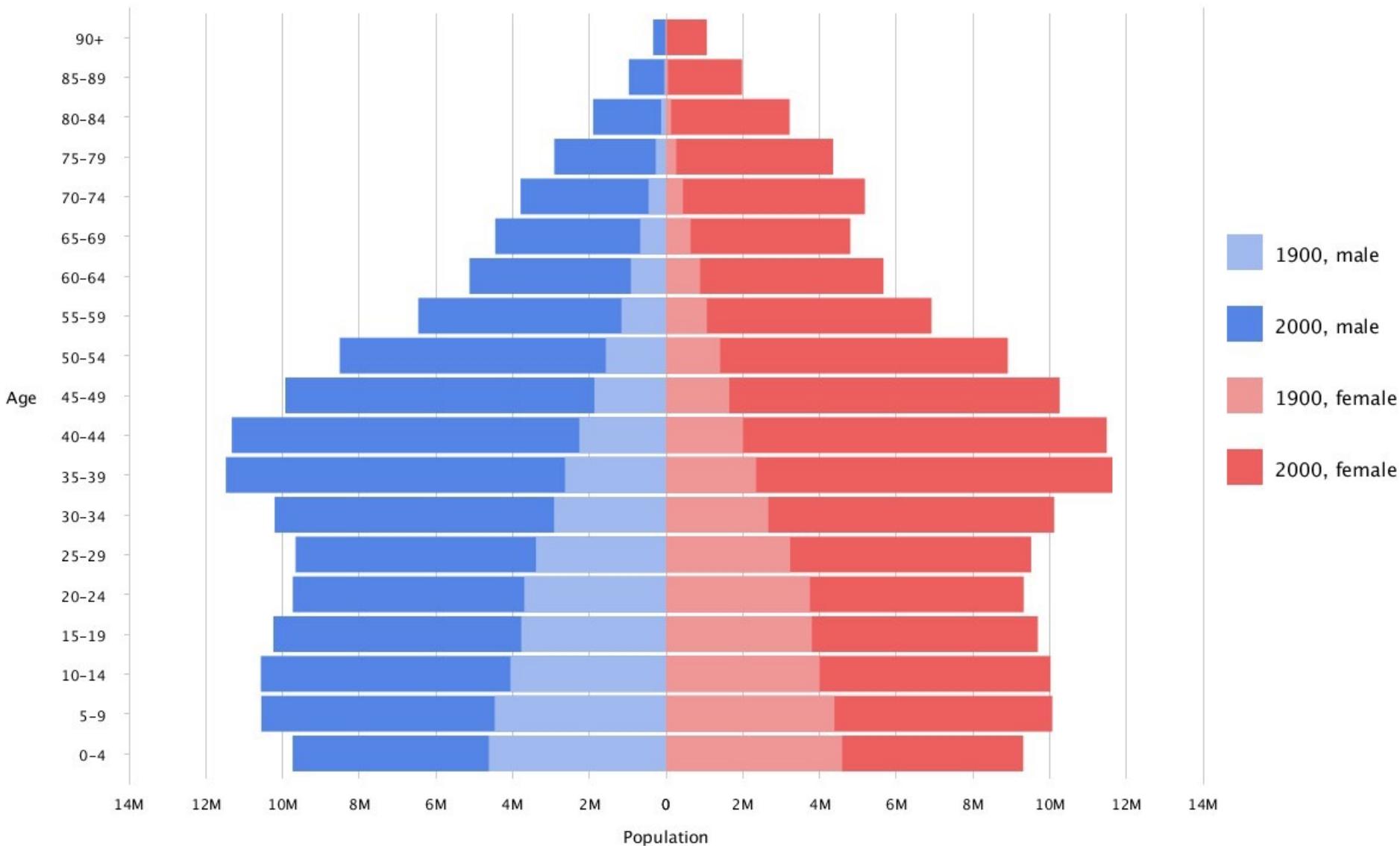


How does the burden of care per adult compare between 1900 and 2000?

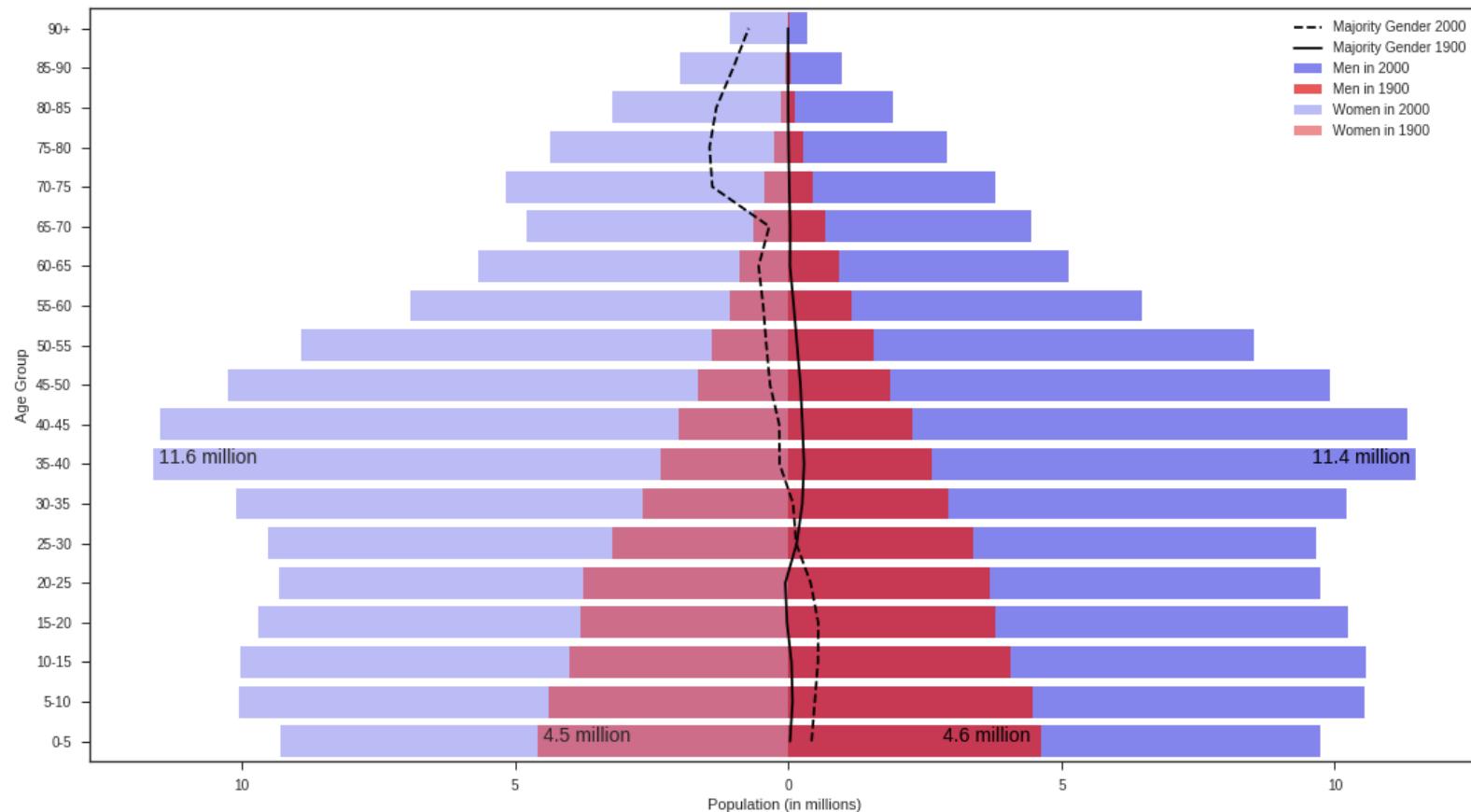


Population Pyramids

How U.S. population's construction changed between 1900 and 2000

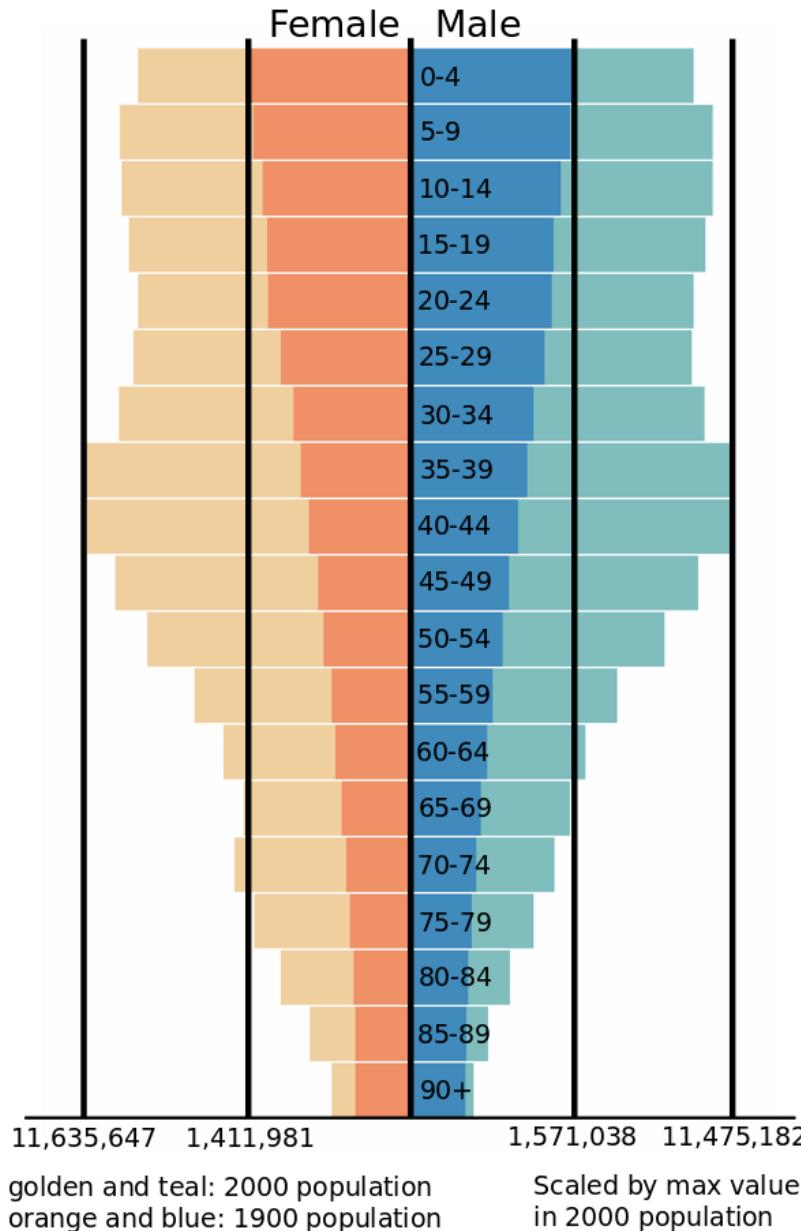


How has the distribution of the US population changed between 1900 and 2000?

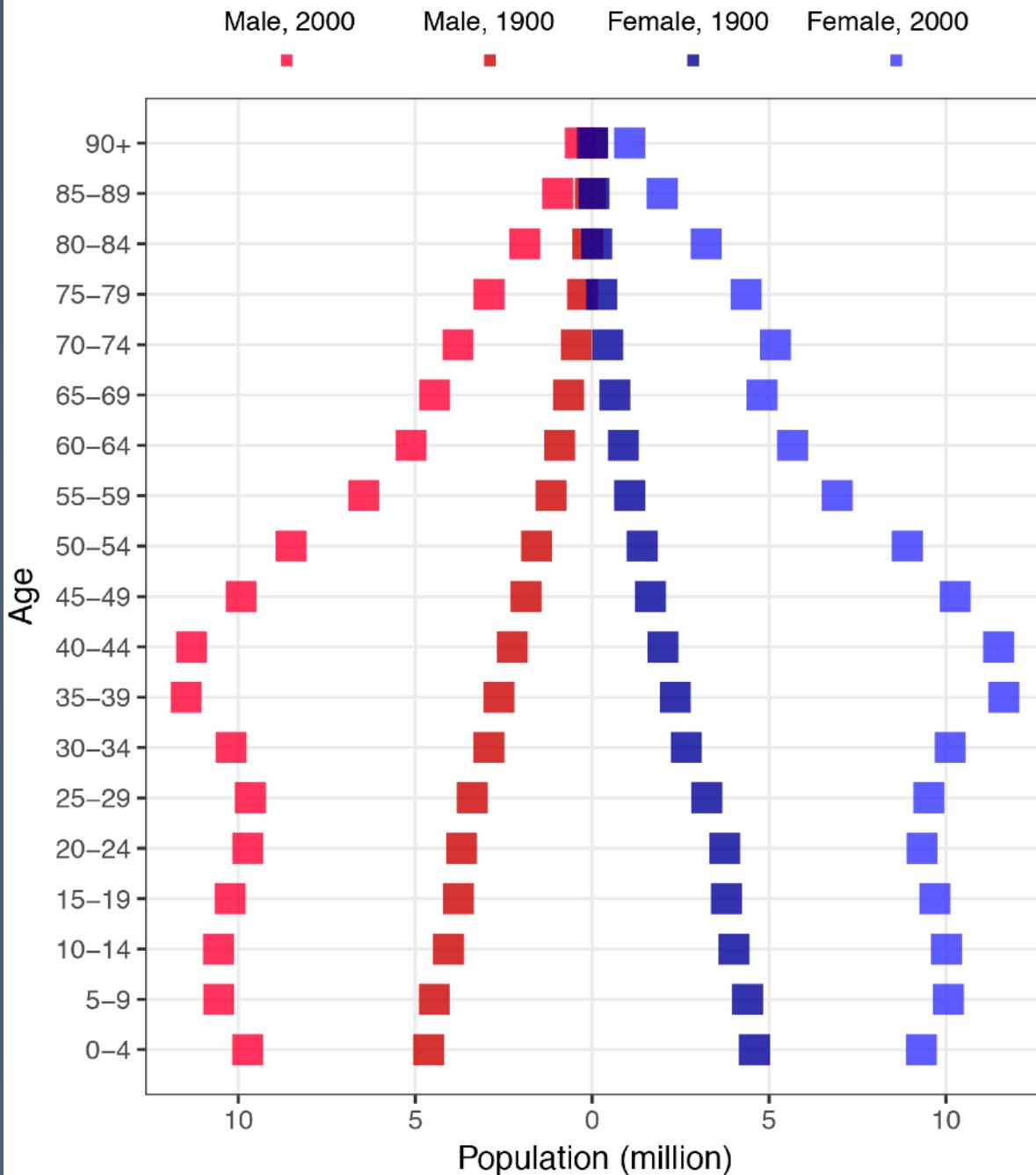


US Population from 1900 to 2000

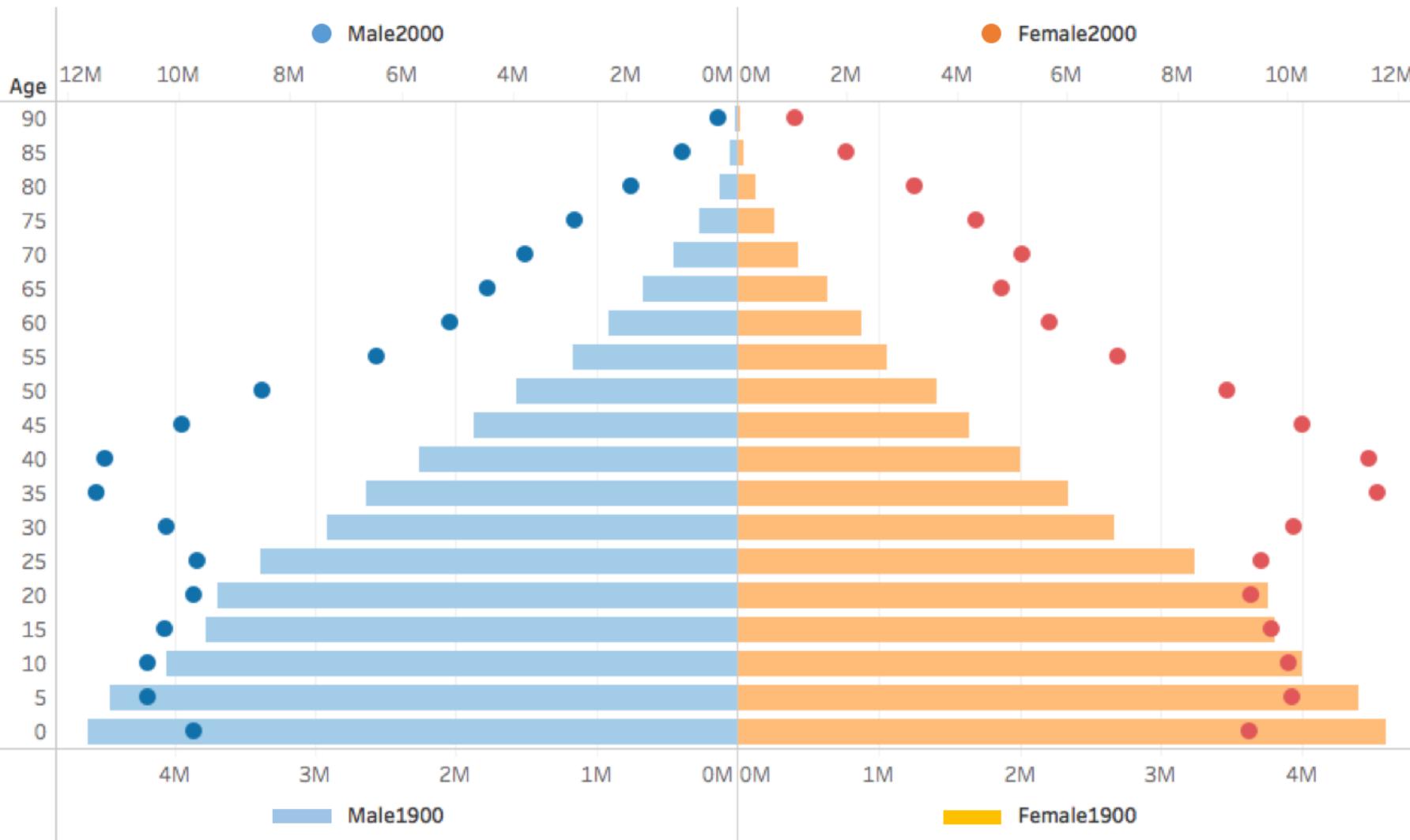
A changing population profile



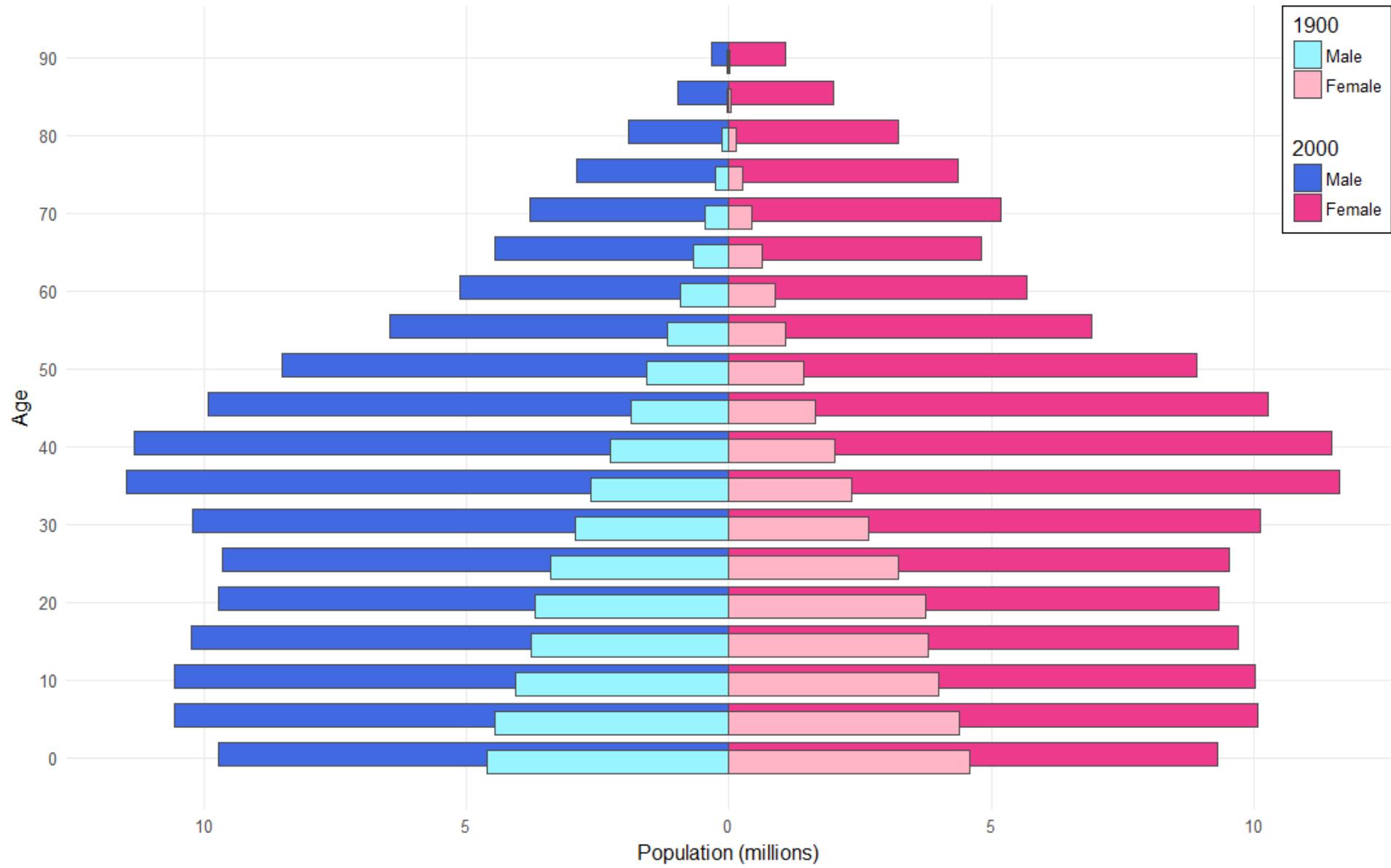
Population Structures in 1900 and 2000 in the USA



Population Structure Change from 1900 to 2000

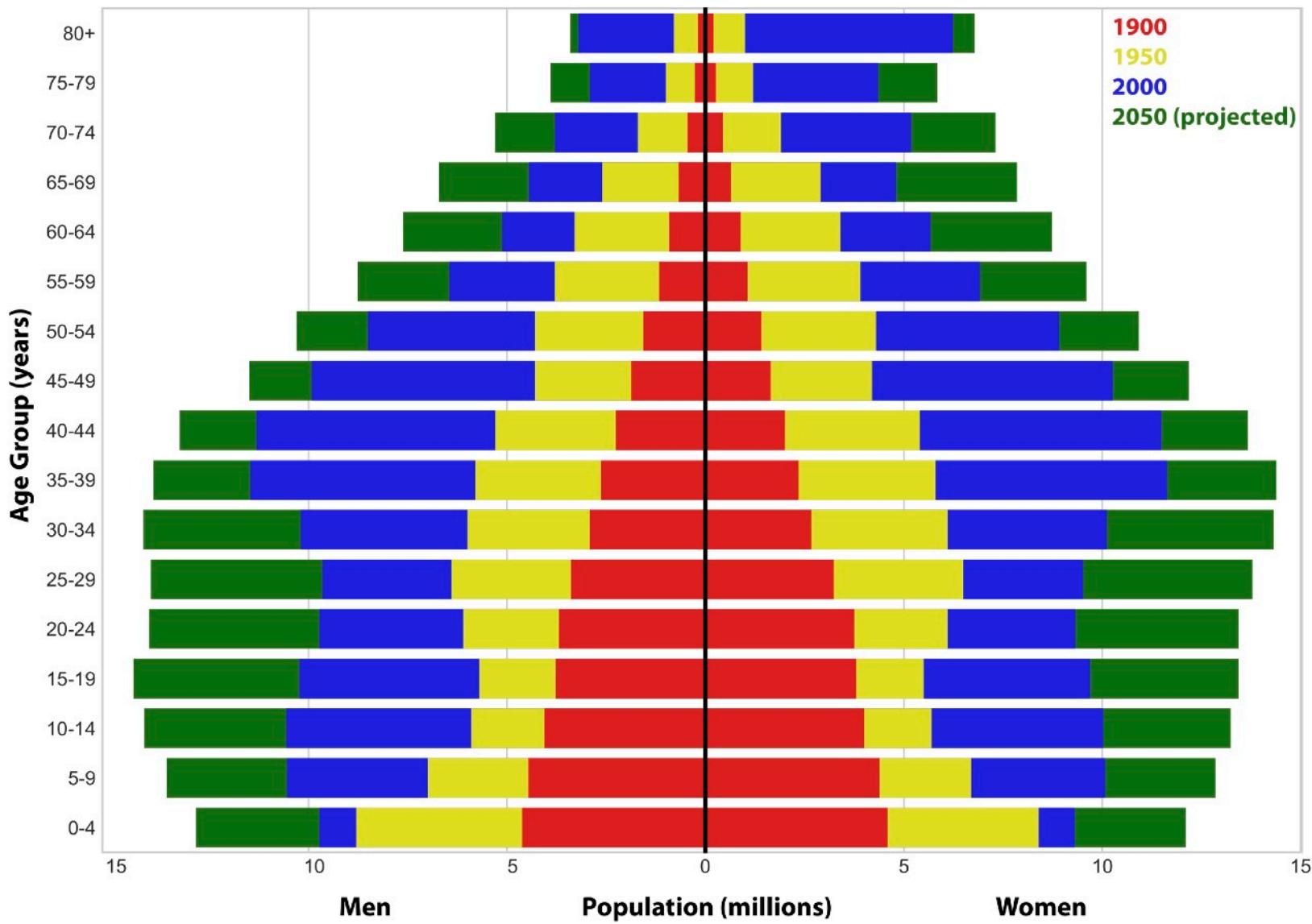


How did the population distribution of the United States change between 1900 and 2000?

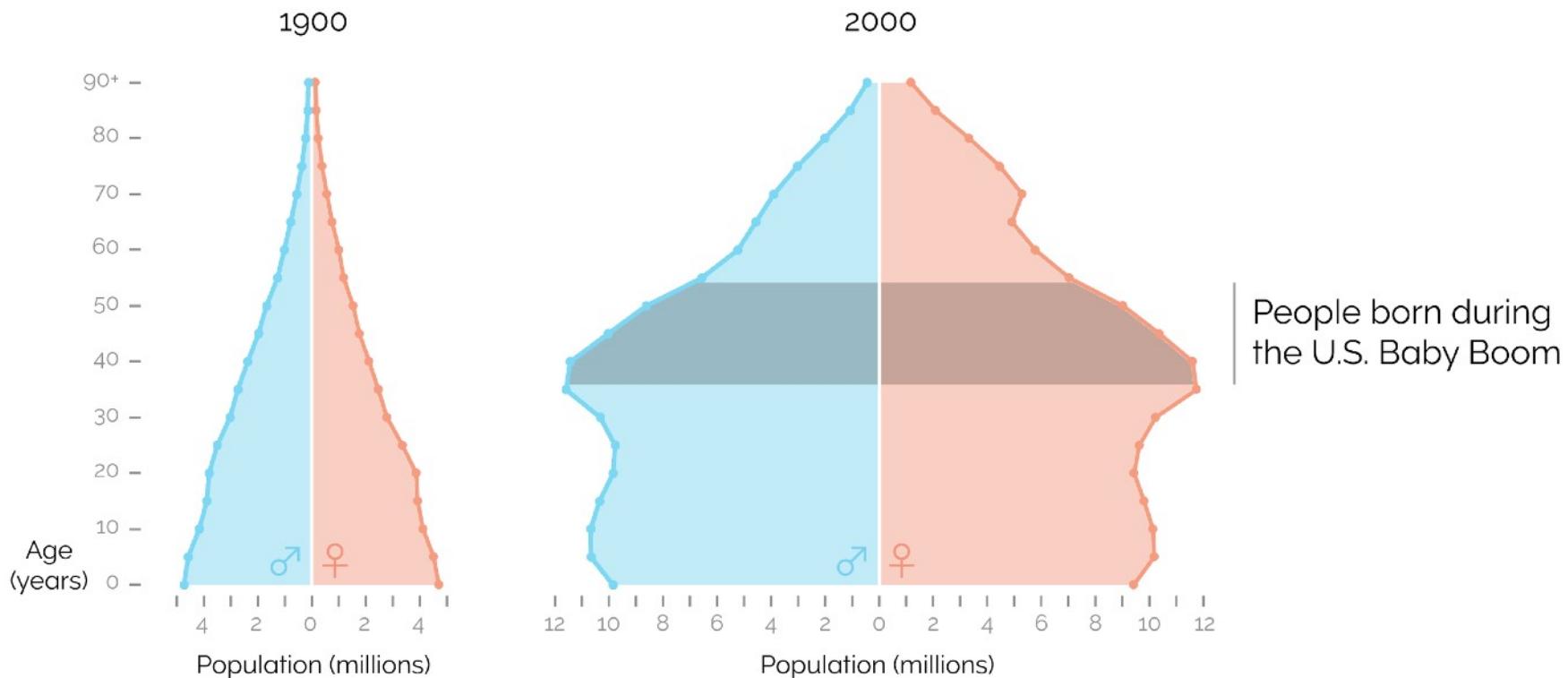


Source: U.S. Census Bureau via IPUMS

What will the US population look like in 2050?

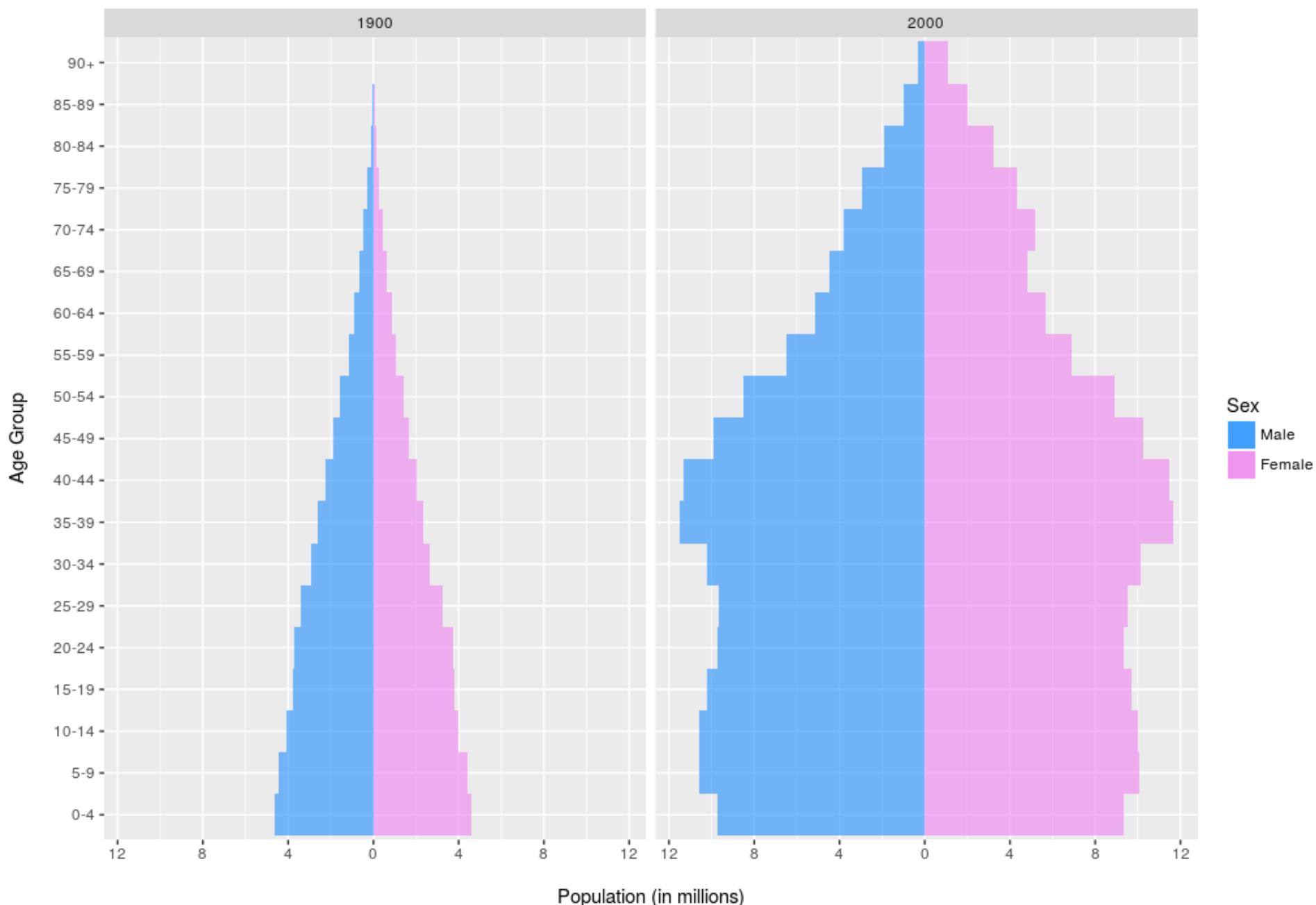


The Impact of Baby Boomers on American demographics

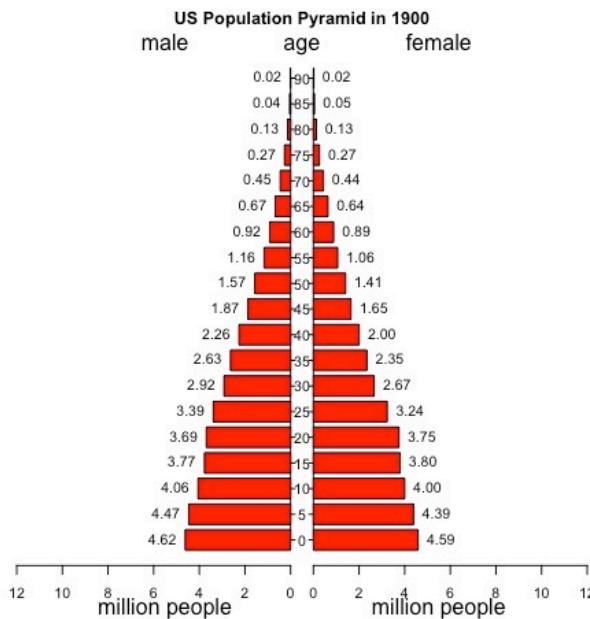
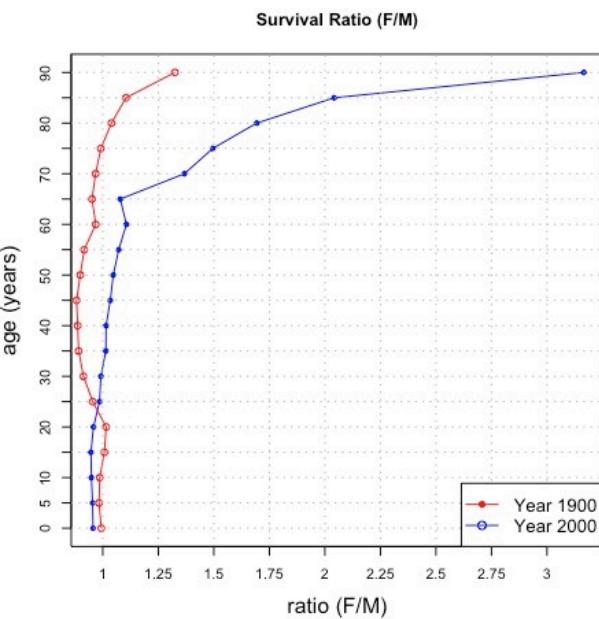


How has the U.S. Population Makeup Changed over a Century?

Population Pyramids for the Years 1900 and 2000

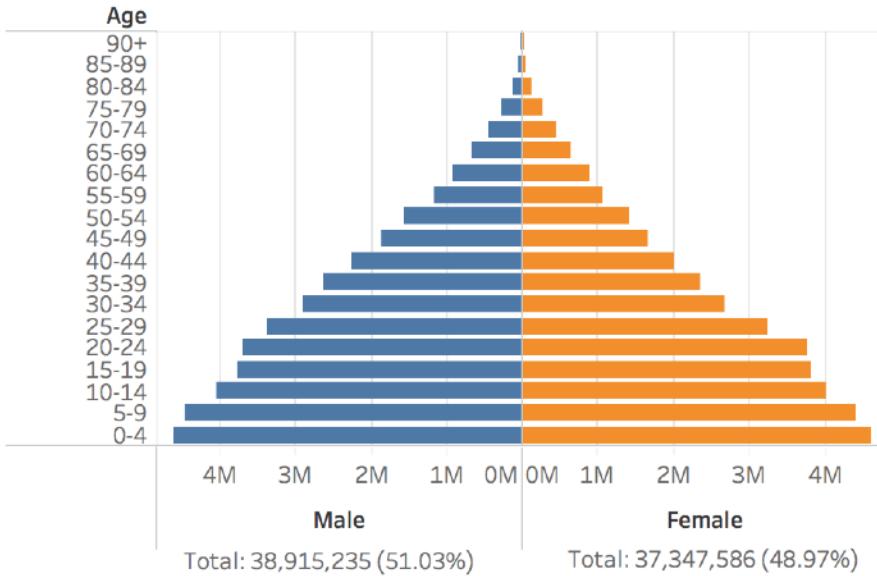


Gender Ratio by Age Brackets in 1900 and 2000

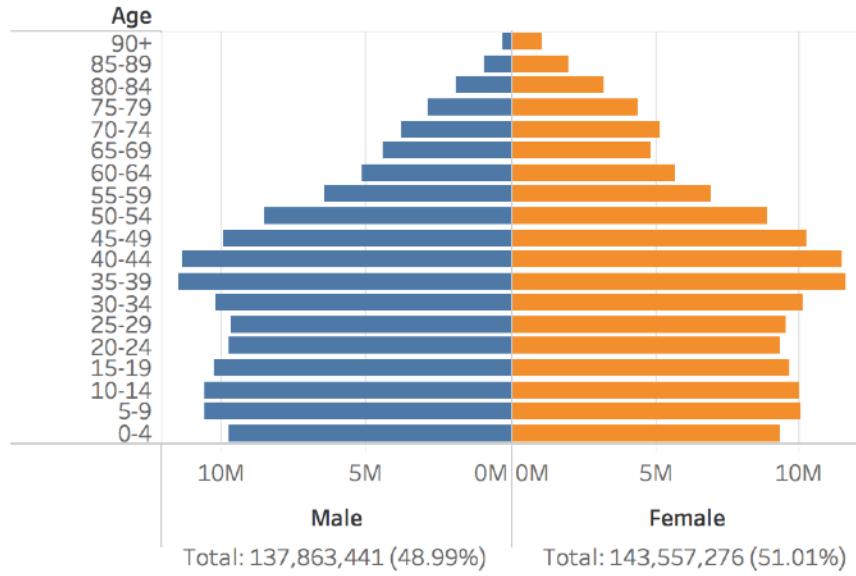


How did U.S. population change in terms of sex and age groups during the 20th century?

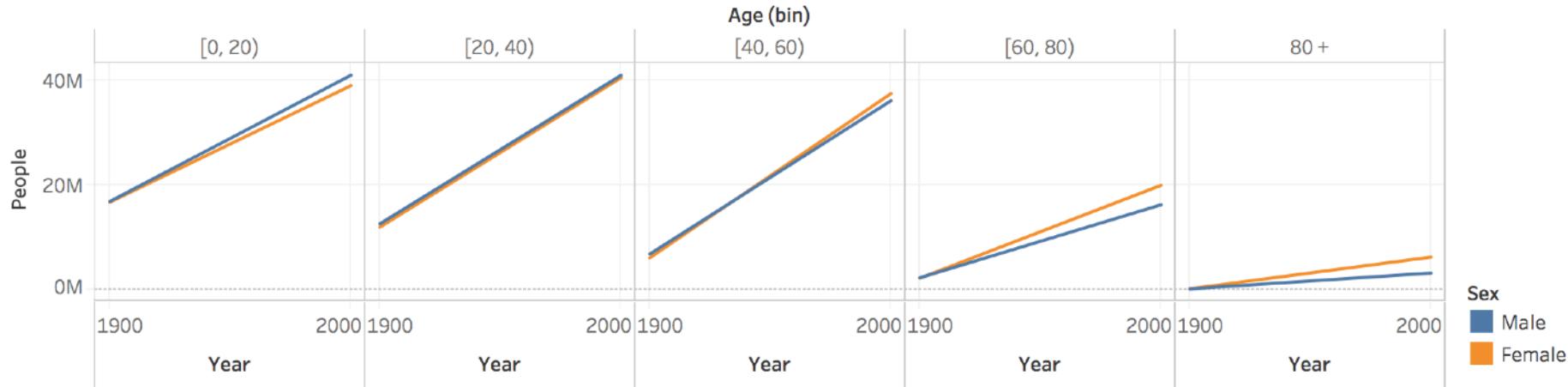
U.S. Population Pyramid in 1900



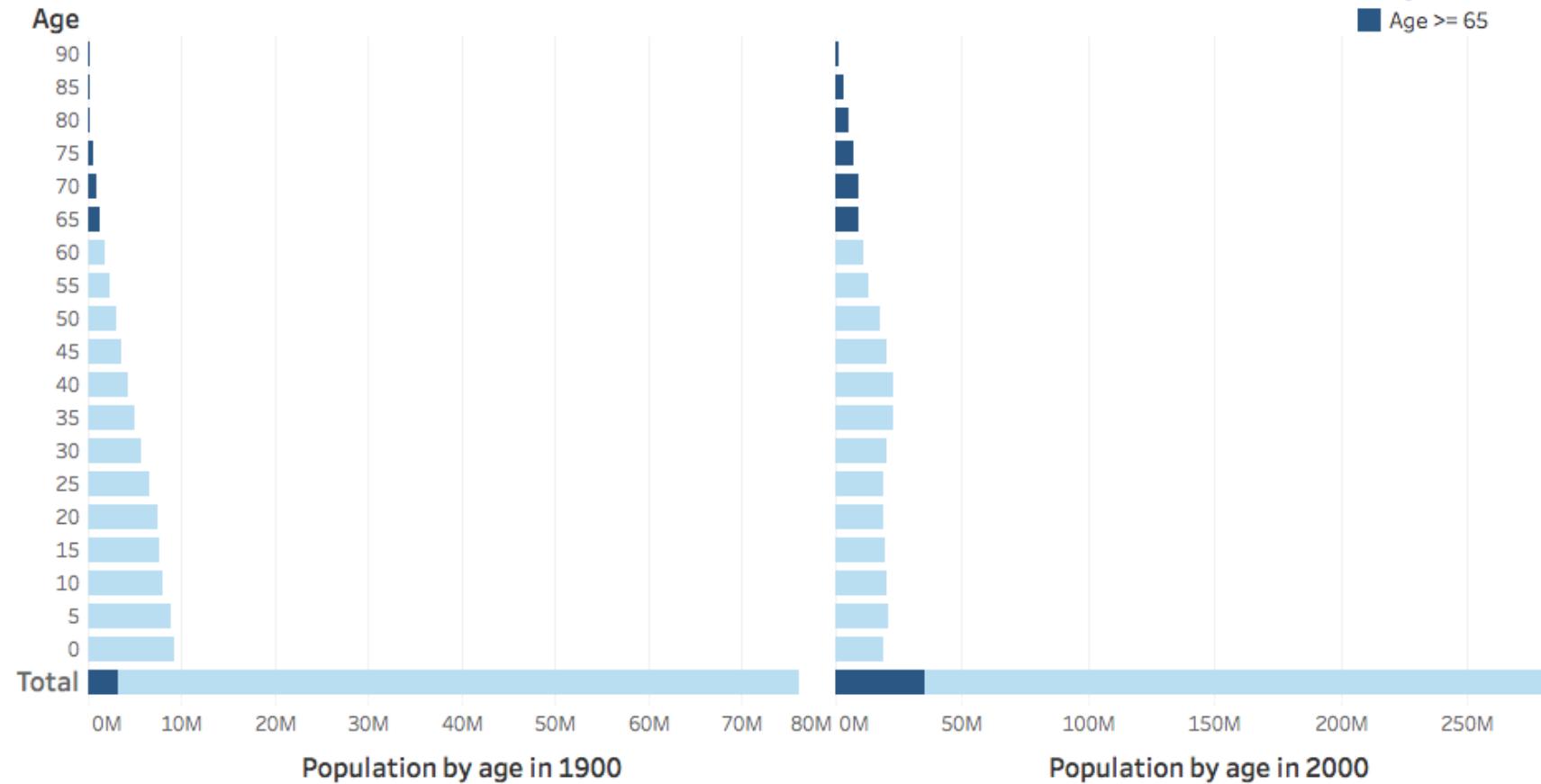
U.S. Population Pyramid in 2000



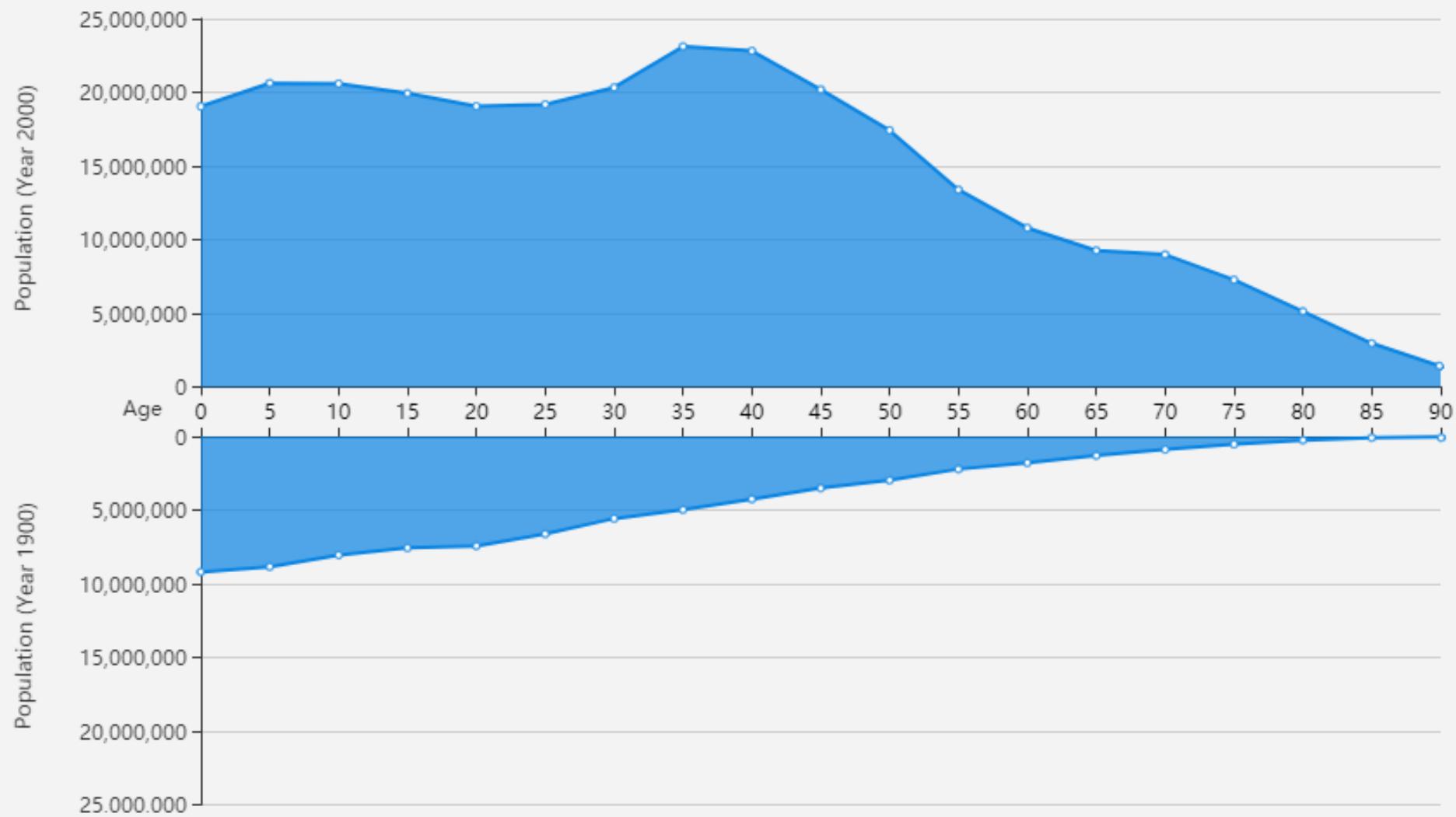
U.S. Population growth trends by sex in different age groups during the 20th century



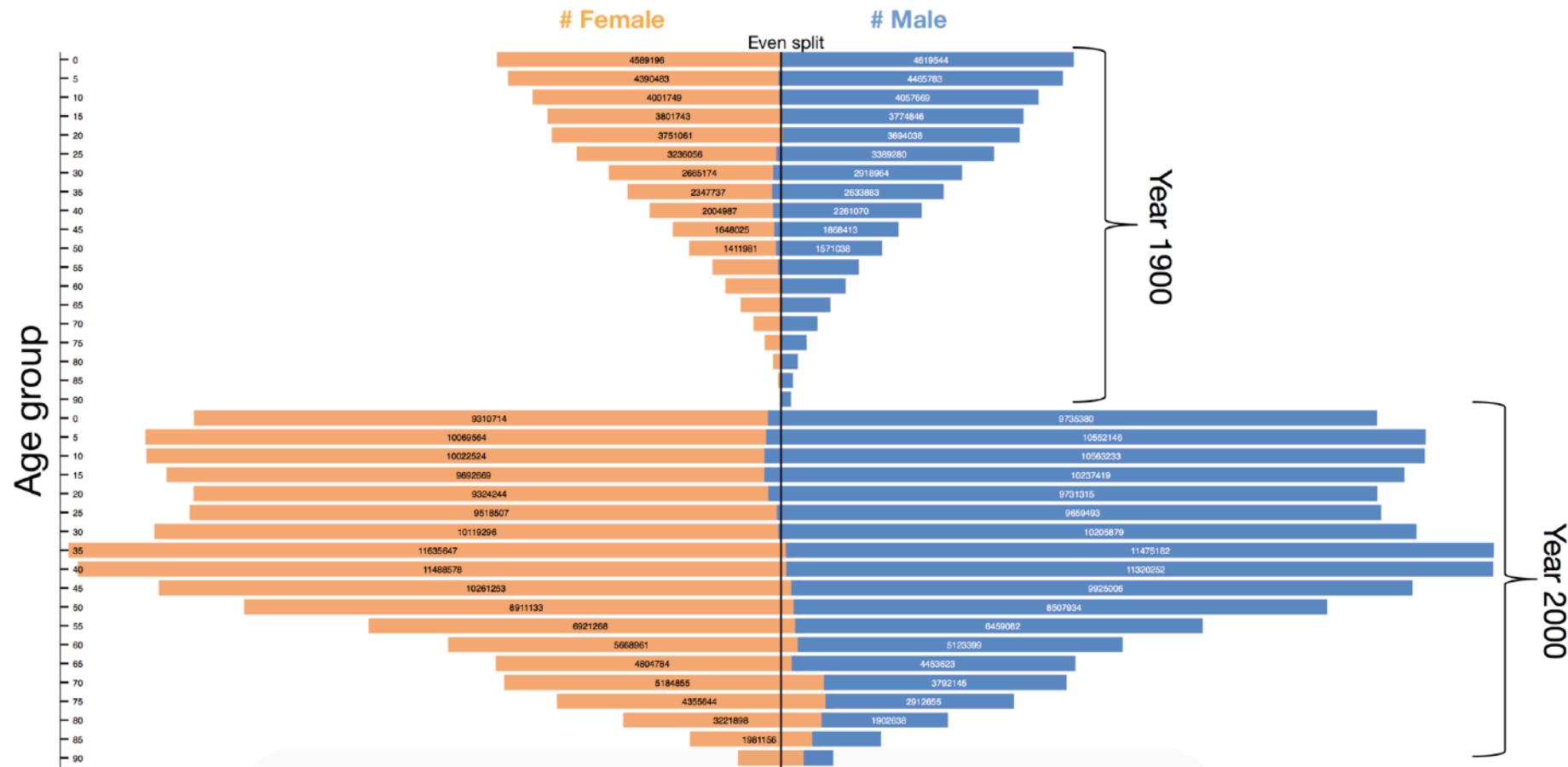
From 1900 to 2000, is the U.S. getting older?



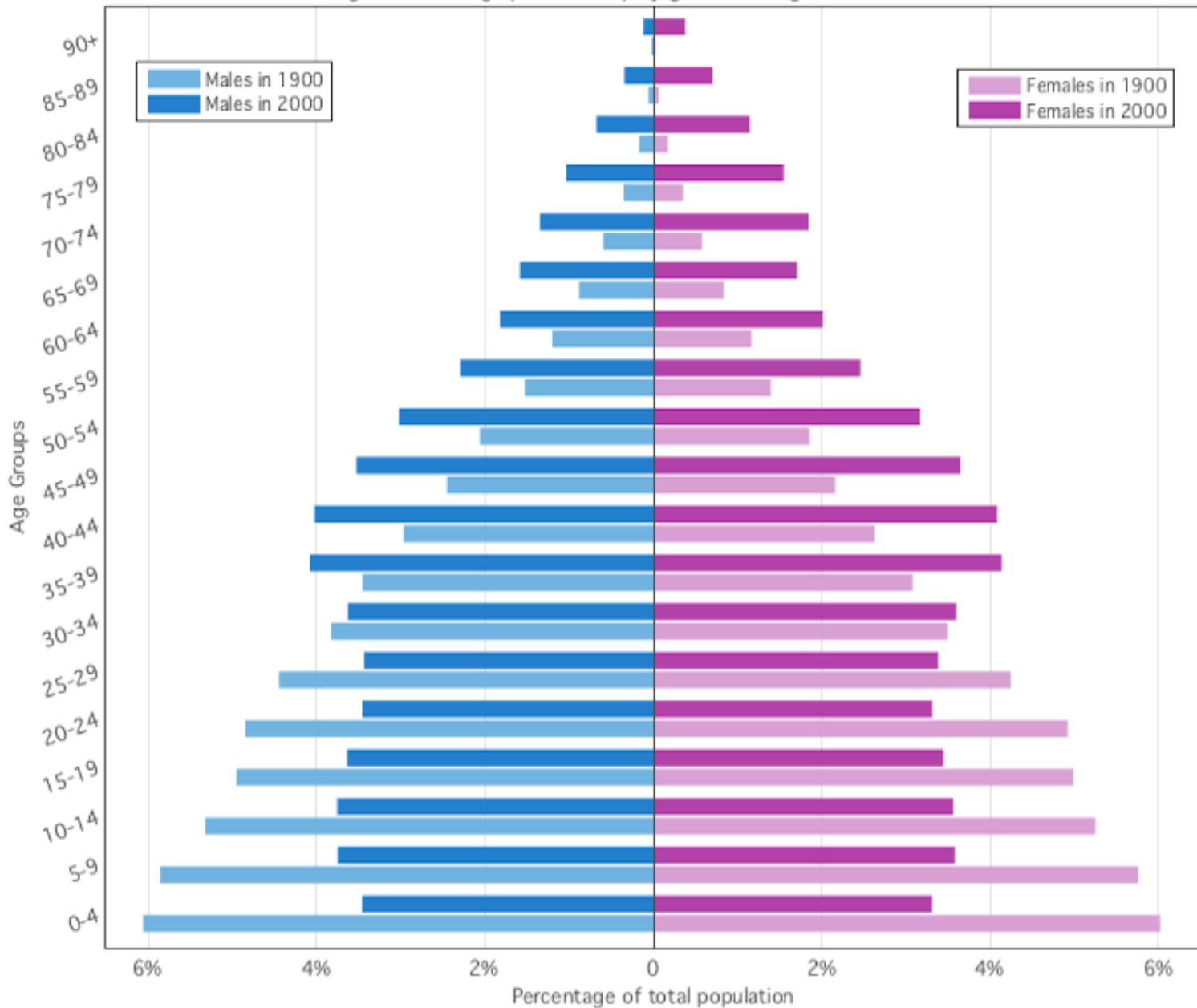
United States Population by Age (Year 2000 vs Year 1900)



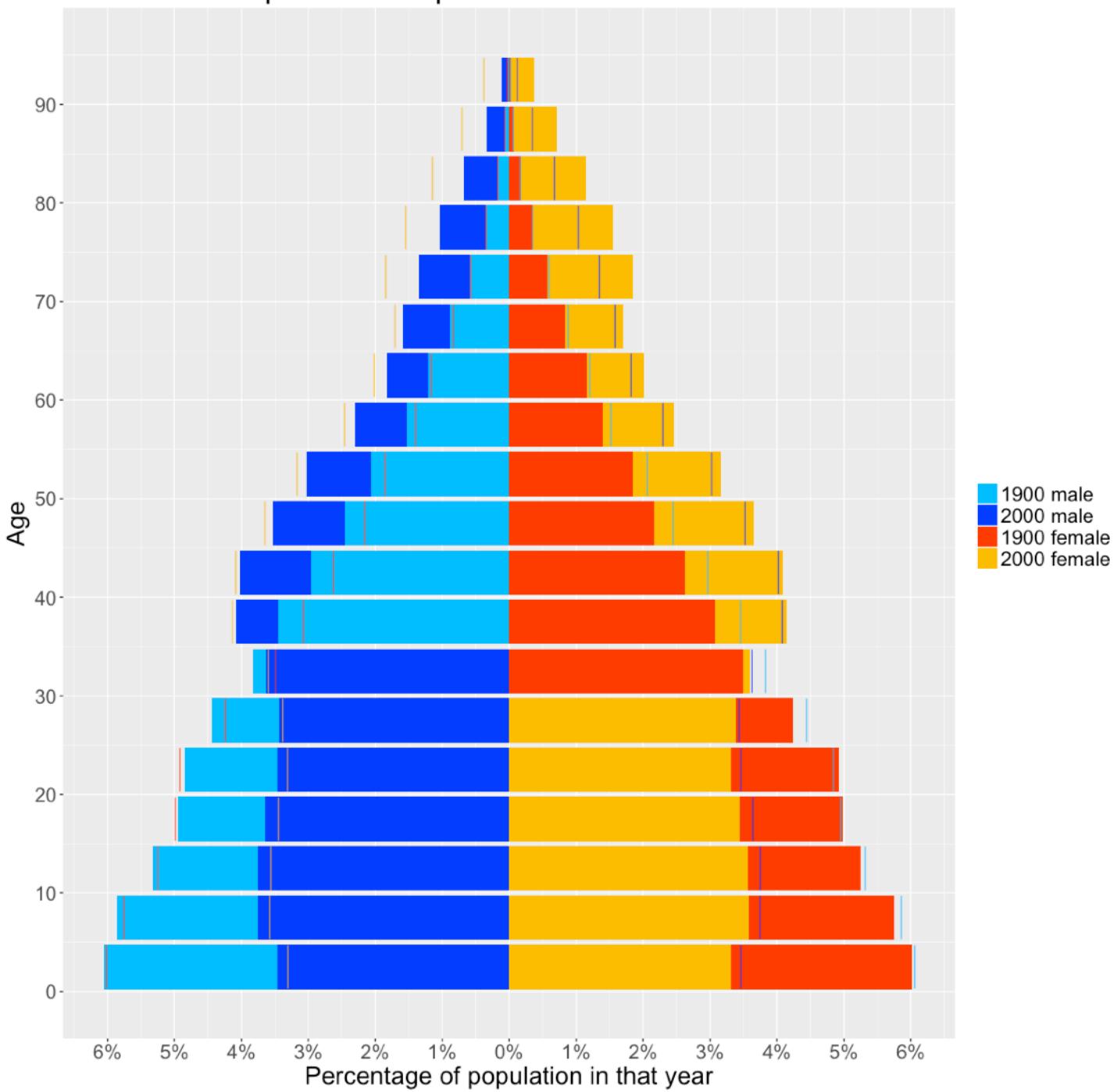
How has the number of women compared to the number of men in the US changed between 1900 and 2000?



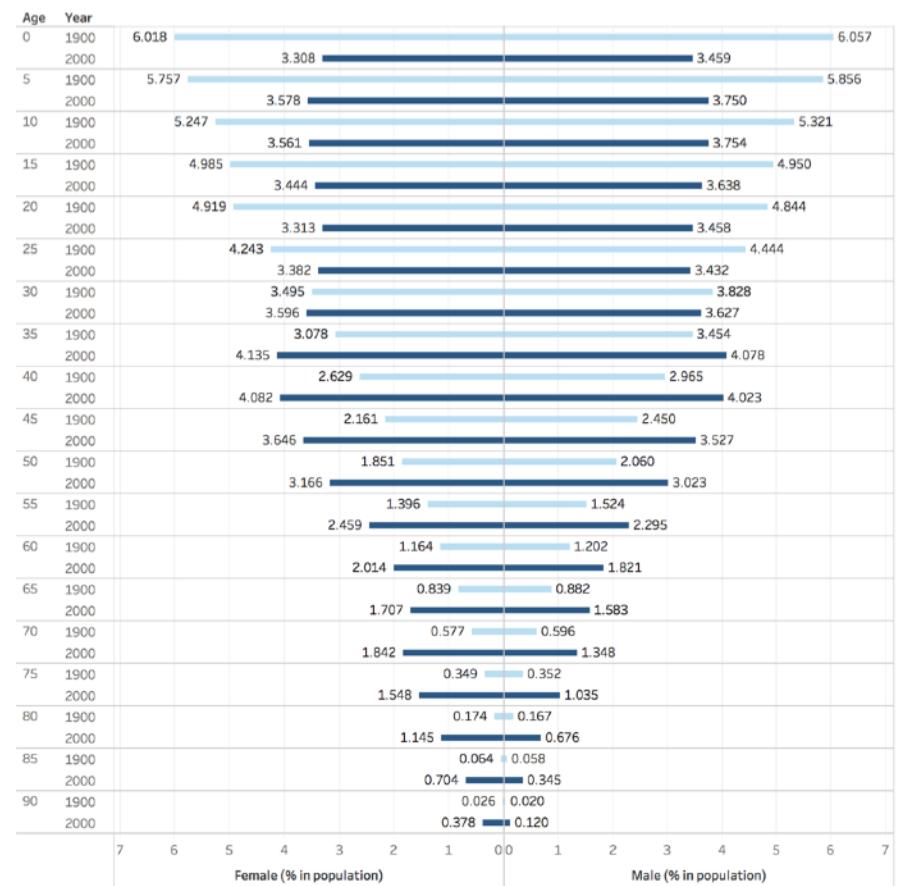
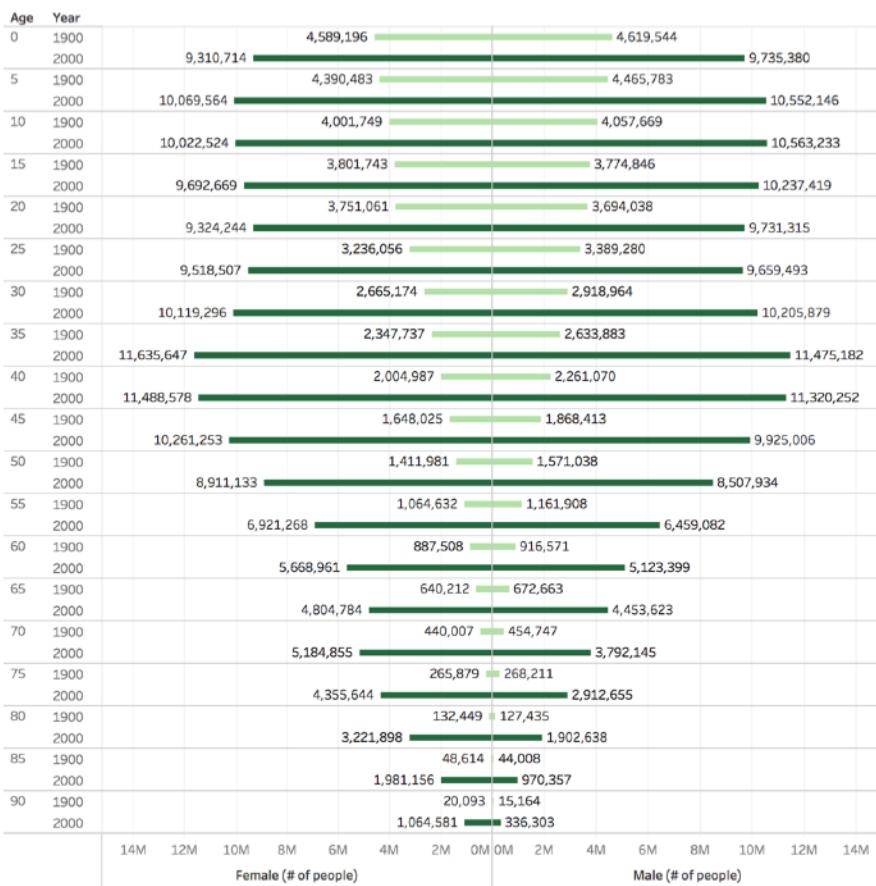
Change in US demographic make-up by gender and age from 1900 to 2000



Population composition for 1900 and 2000



How does the demographic make-up of the United States differ between the two years 1900 and 2000?



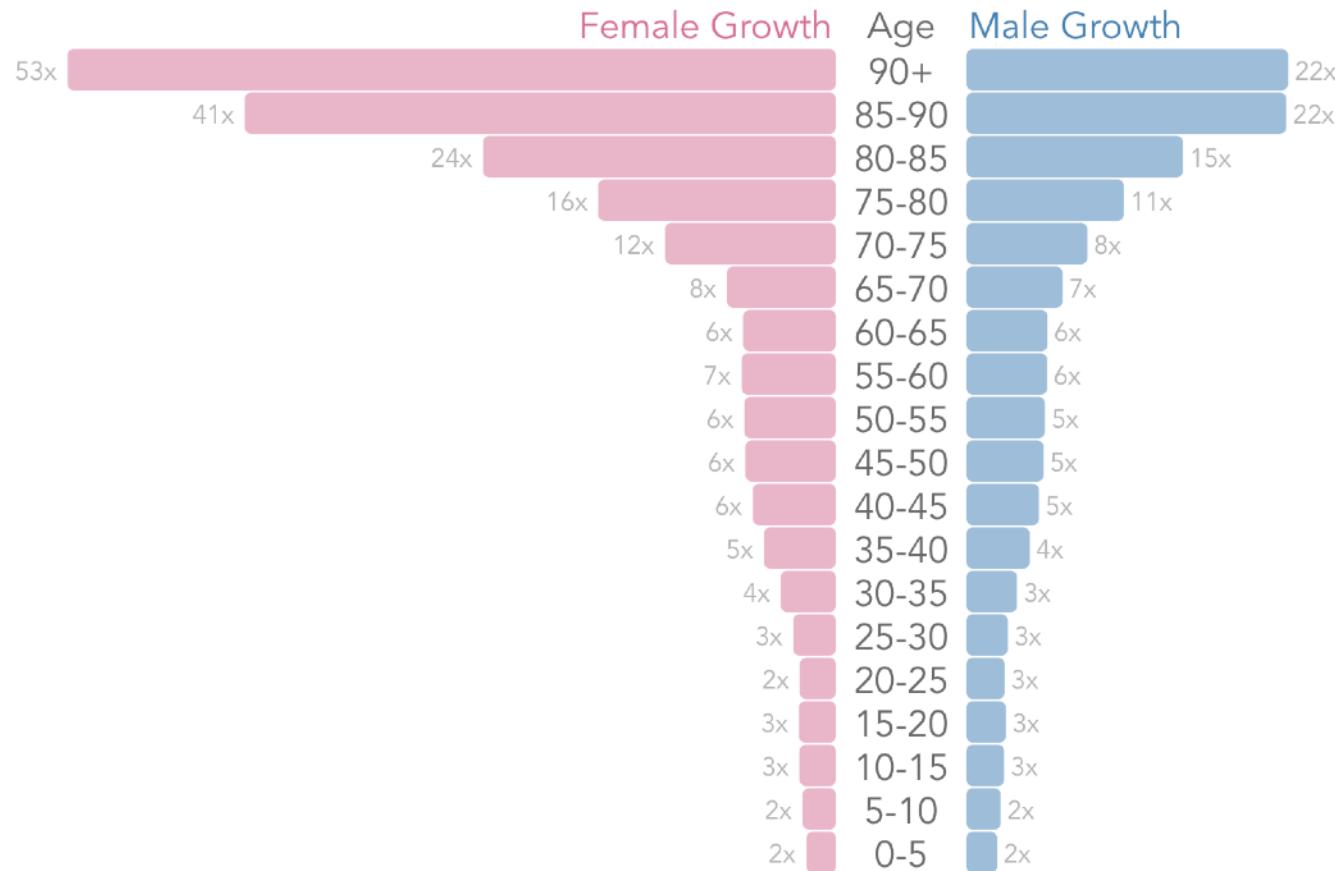
Sum of Female and sum of Male for each Year broken down by Age. Color shows details about Year.



Sum of FemalePct and sum of MalePct for each Year broken down by Age. Color shows details about Year.

Is The Old Lady Apocalypse Upon Us?

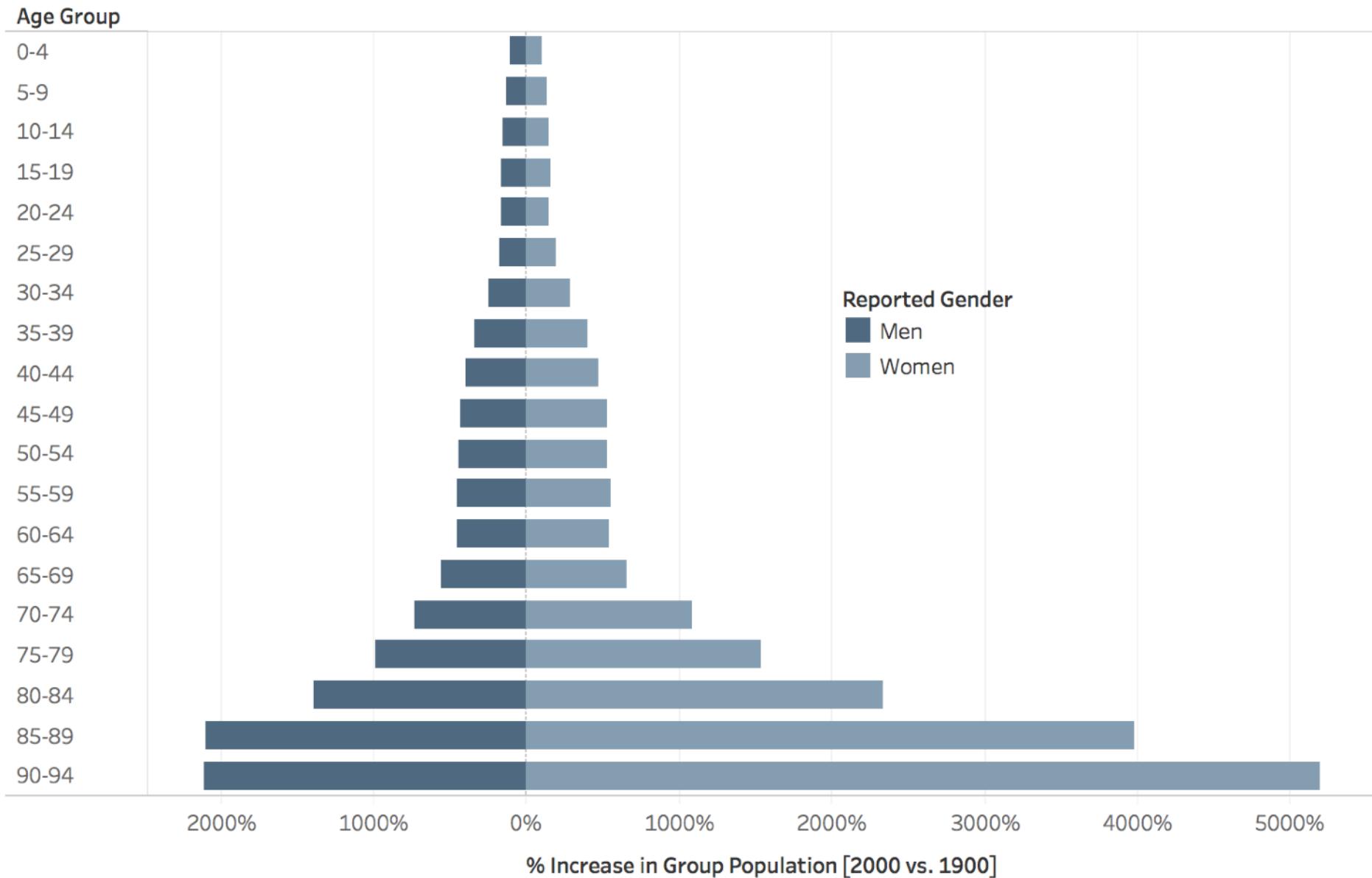
Or which age groups grew the most between 1900 and 2000?



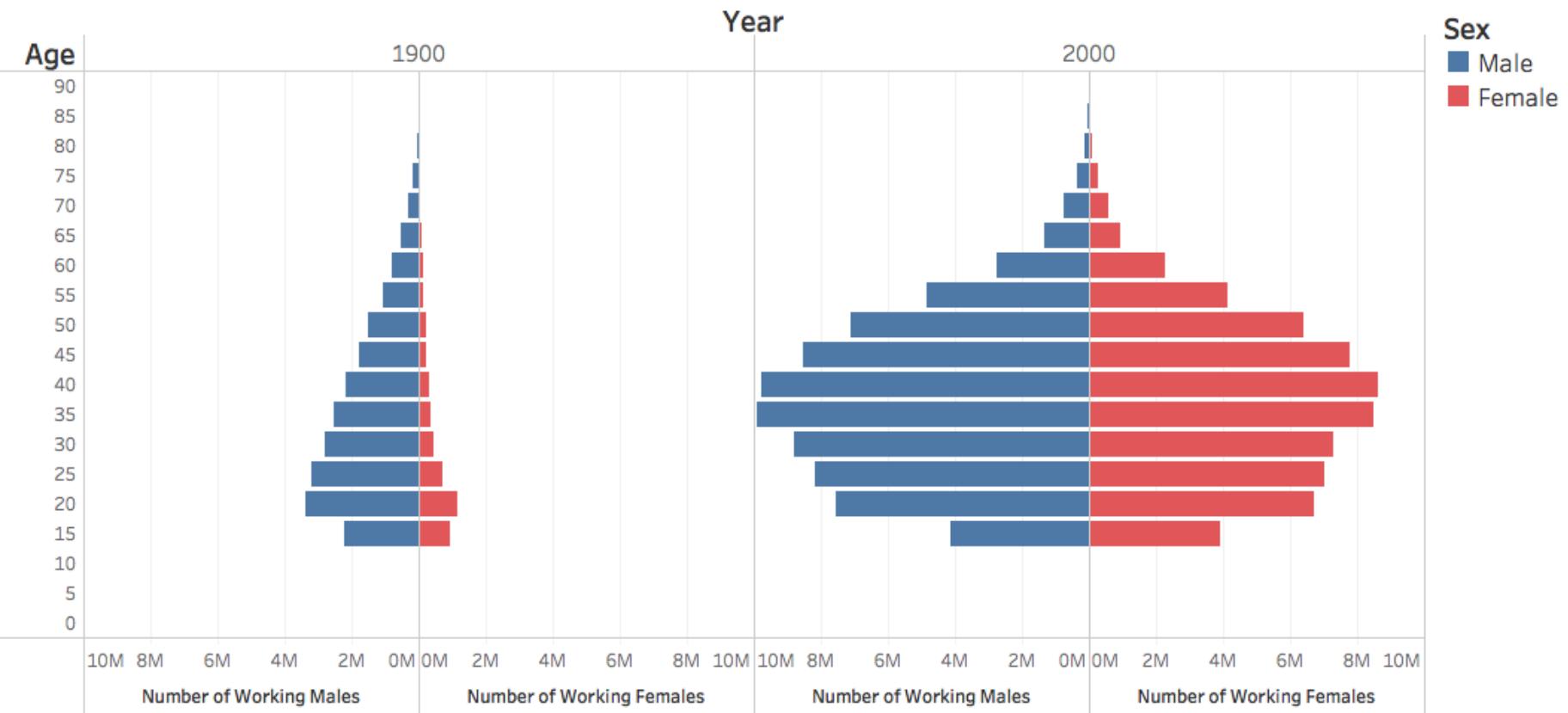
Overall Population Growth: 3.7x (1900-2000)

psb.re/cse512-a1

How have age distributions shifted over the 20th century?

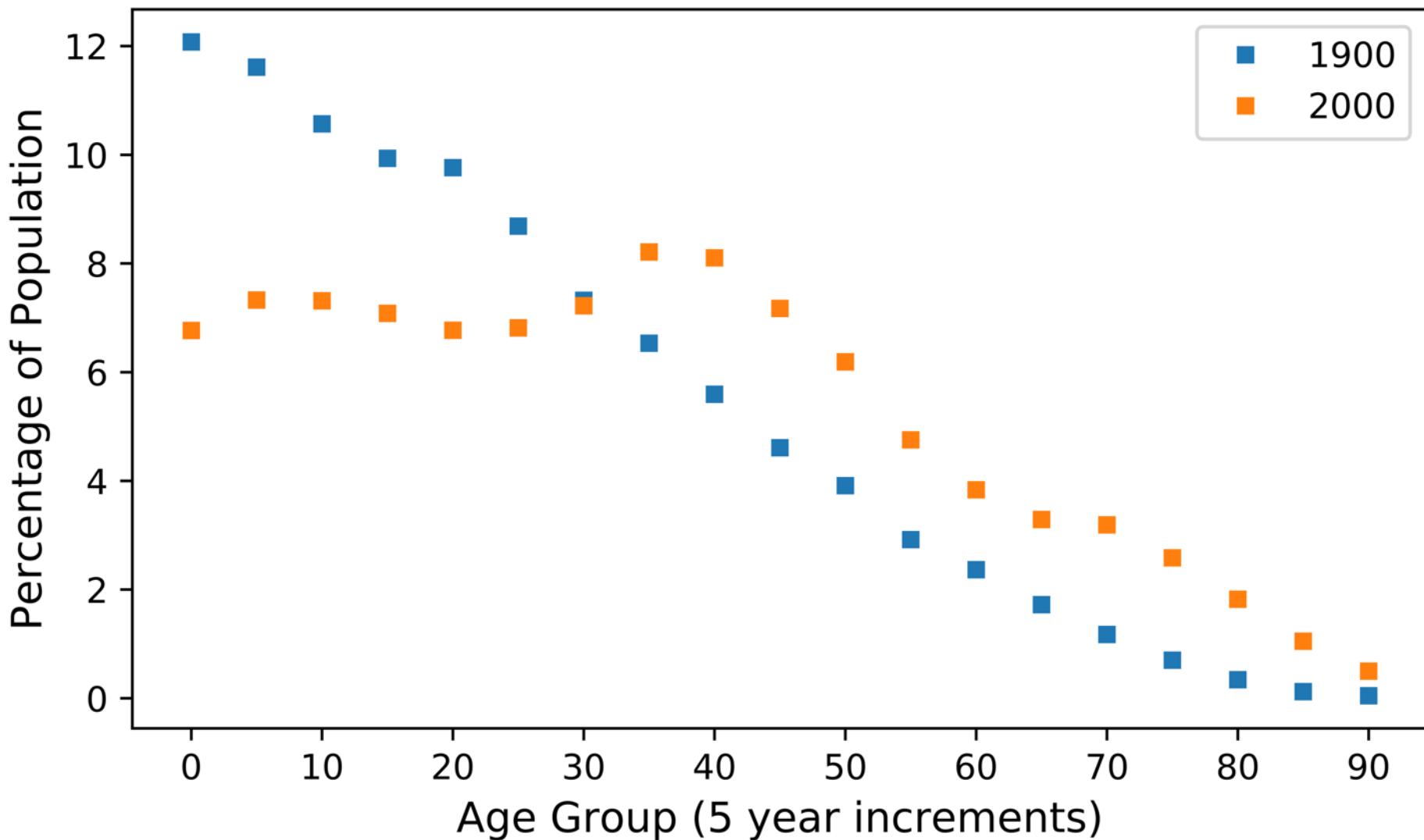


Number of people in the labor force
by
age and sex

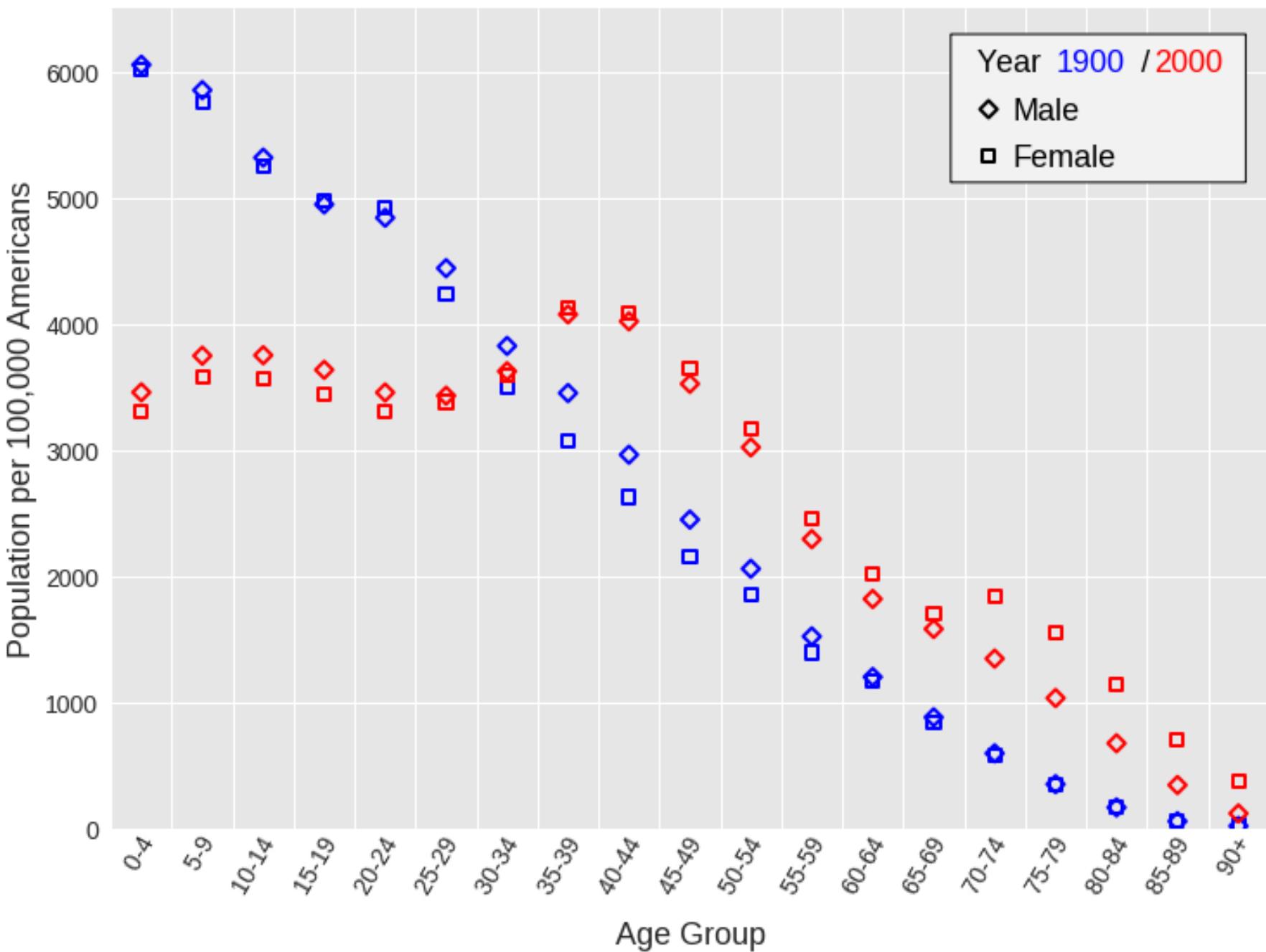


Scatter / Dot Plots

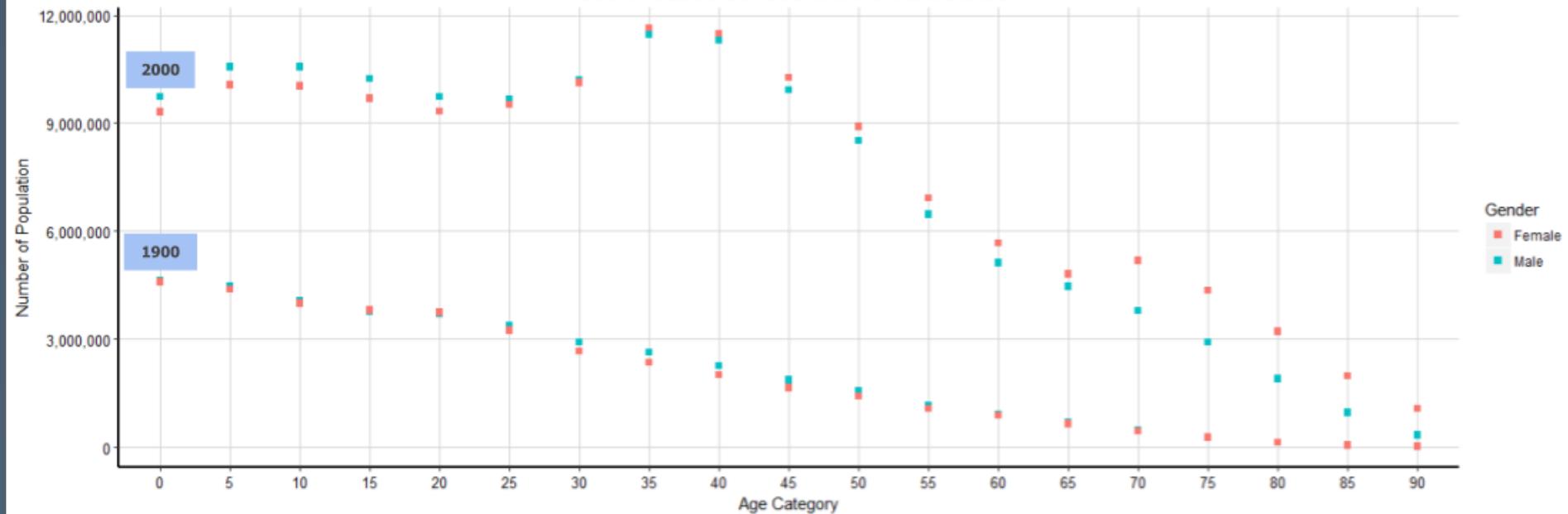
How Did the Age Distribution Change Between the Years 1900 and 2000



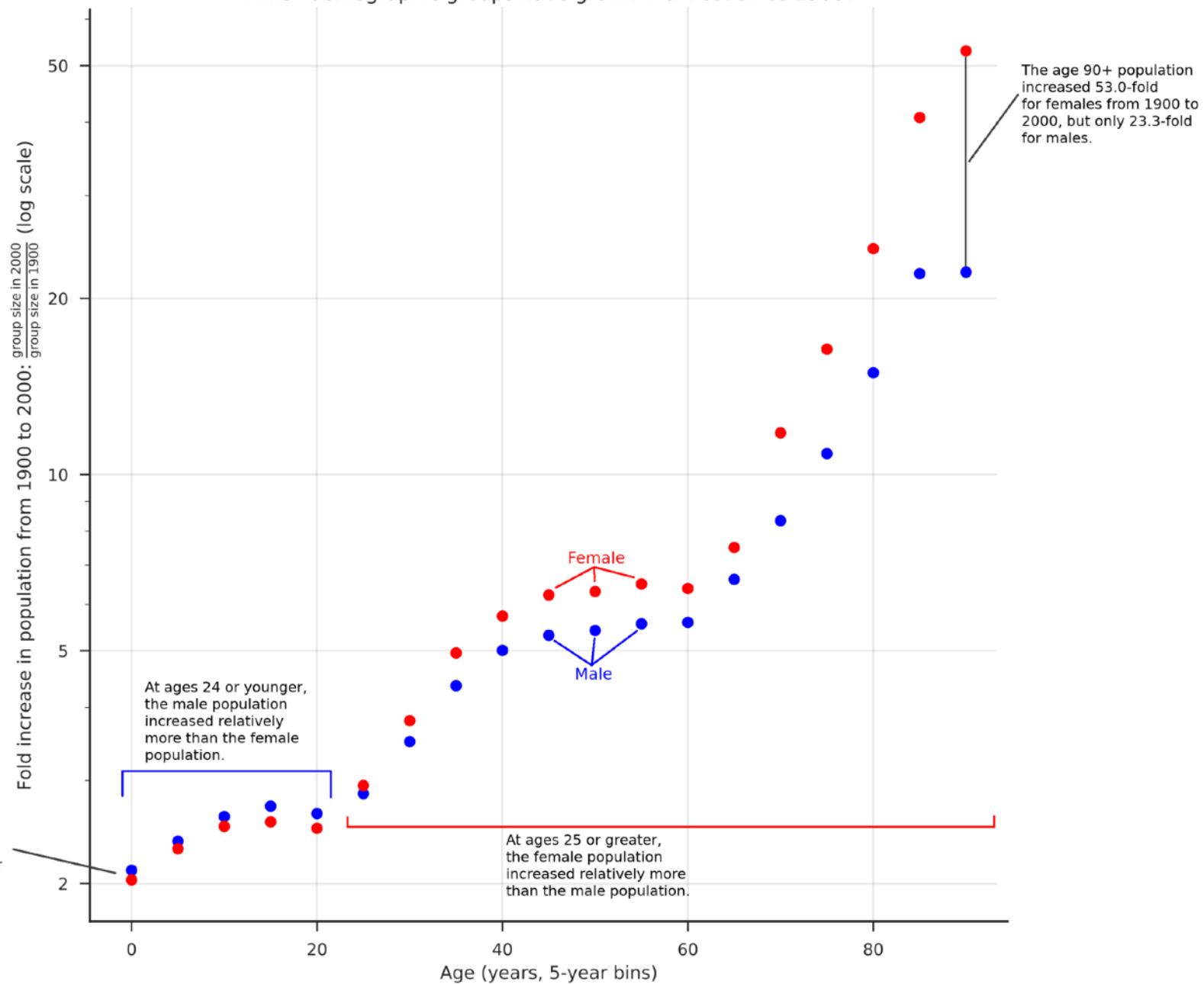
How did the distribution of Americans' age change in the 20th century?



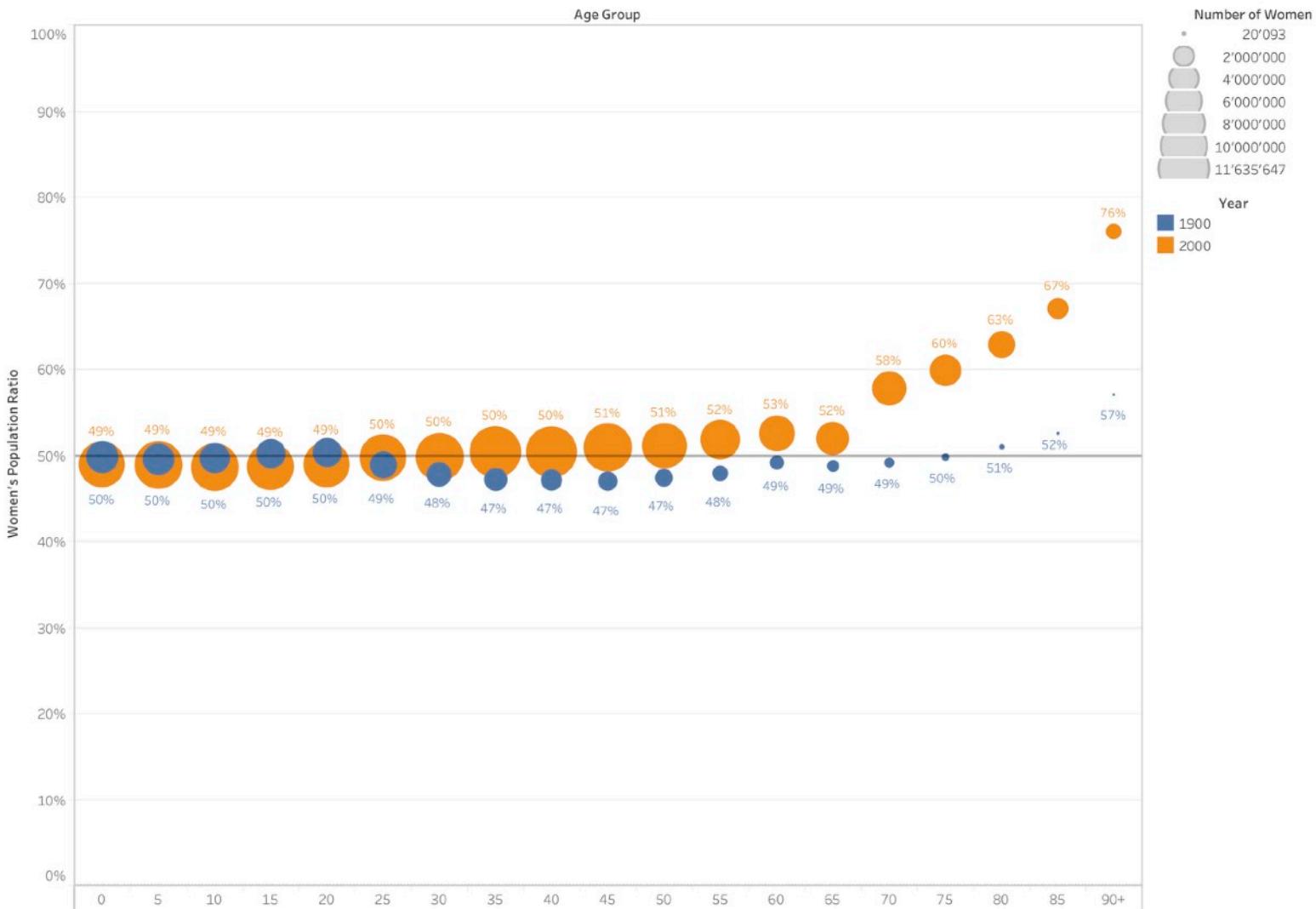
1900 and 2000 Census in the United States



Which demographic groups have grown the most since 1900?



Women's Population Share per Age Group in the US 1900 versus 2000



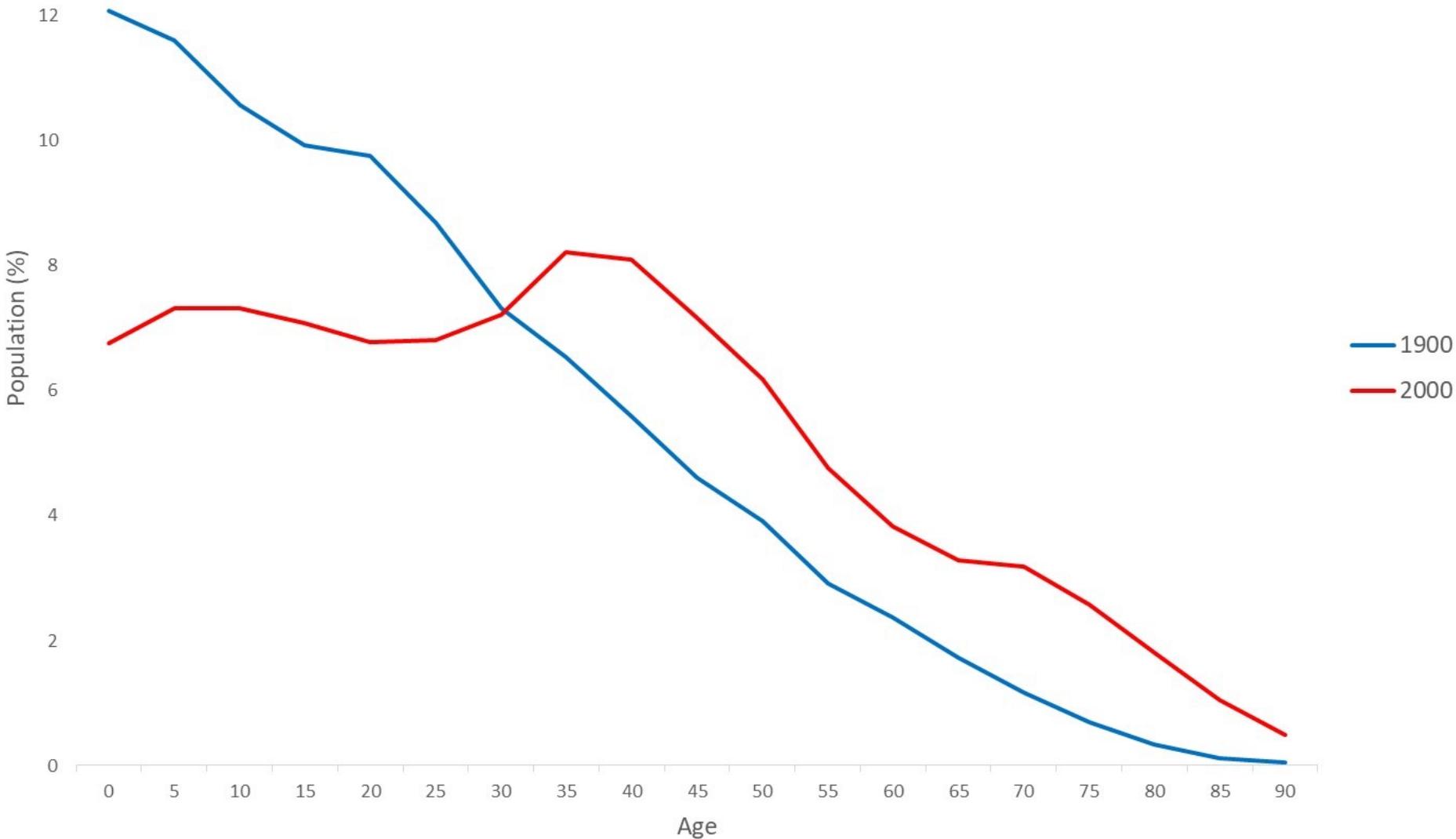
Women's Population Ratio for each Age Group. Color shows details about Year. Size shows total number of women. The marks are labeled by the ratio.

Line & Area Charts

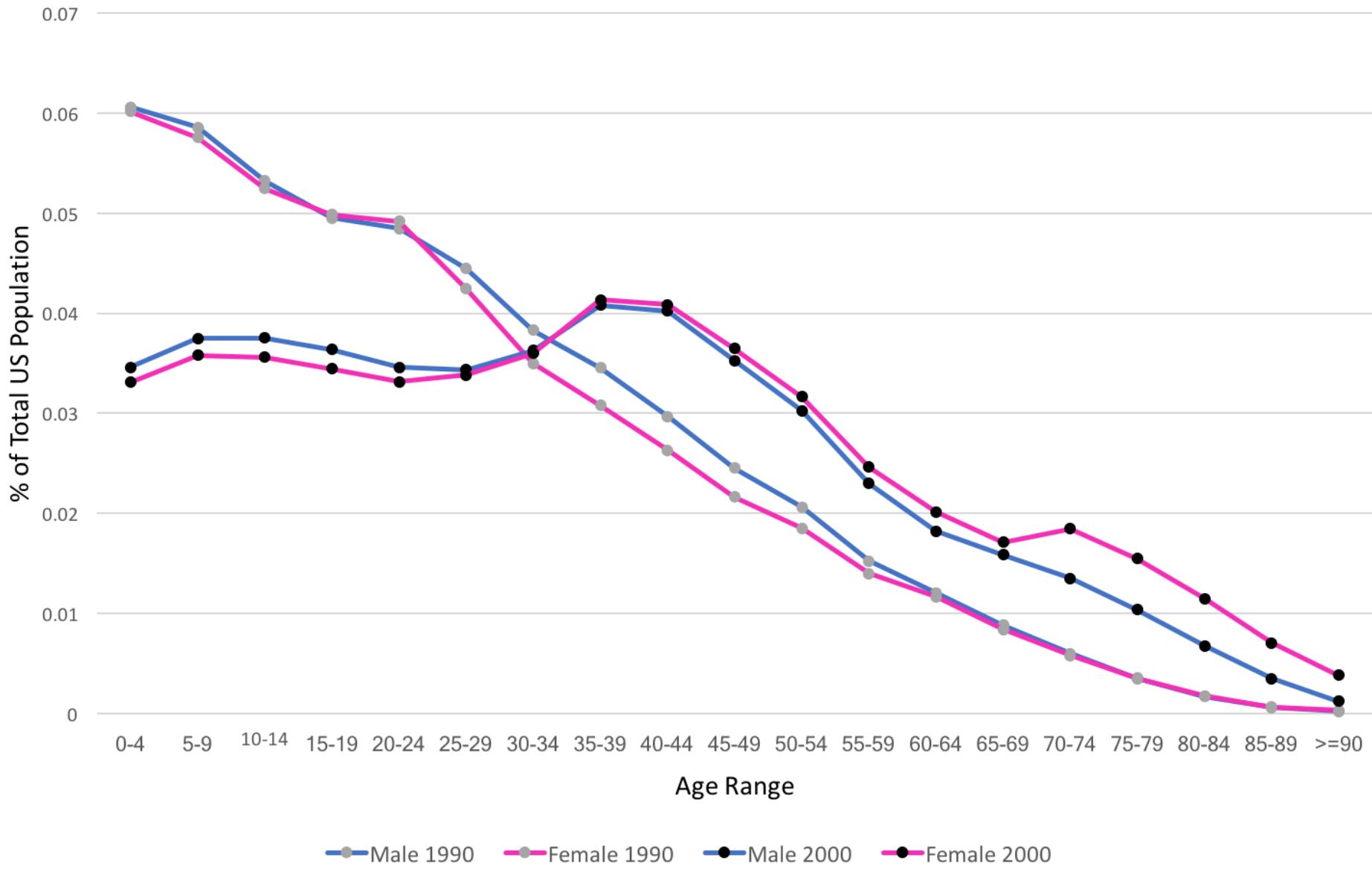
Age distribution of the U.S. population, 1900 vs. 2000

(Source: U.S. Census Bureau via IPUMS)

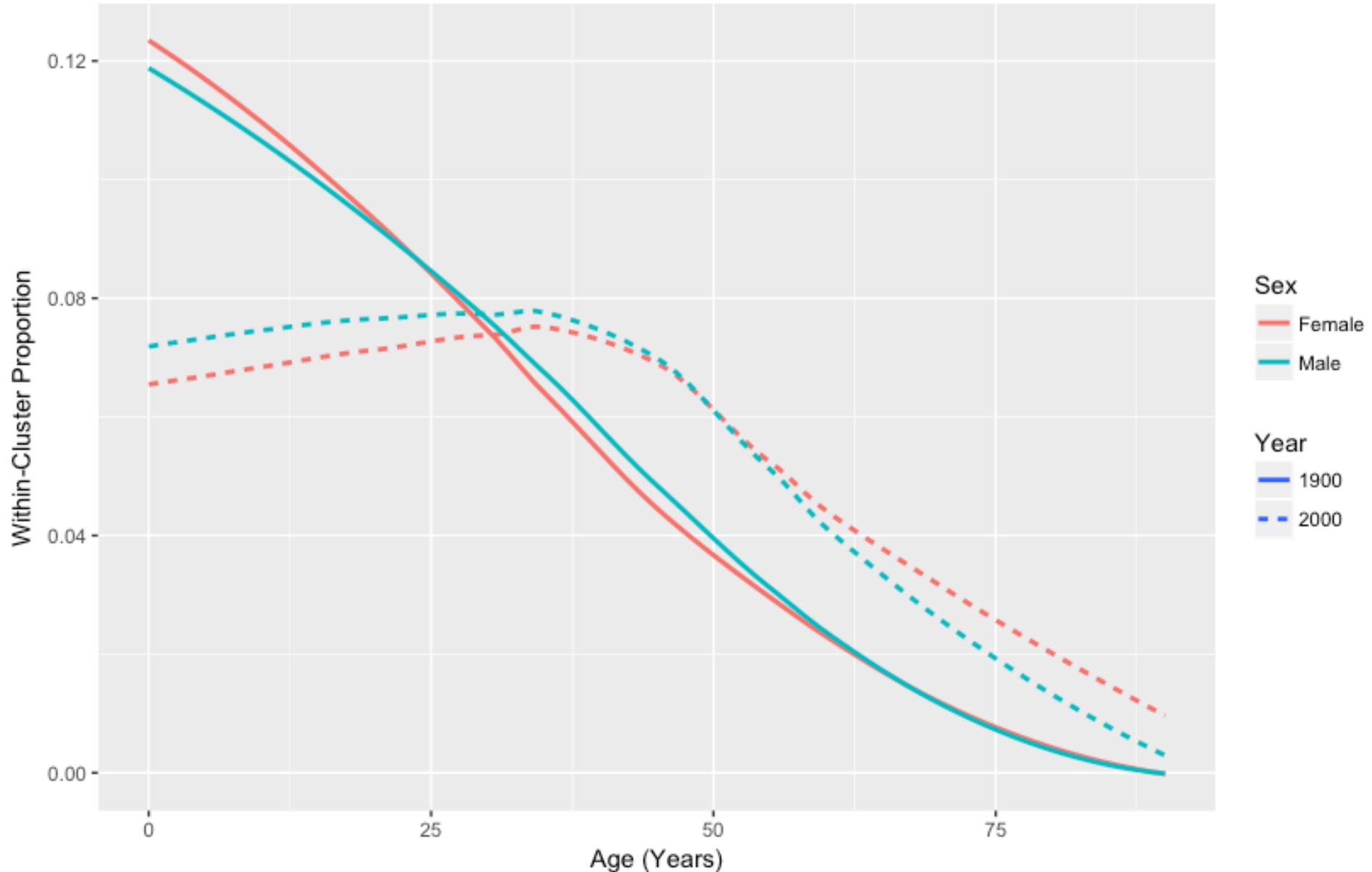
14



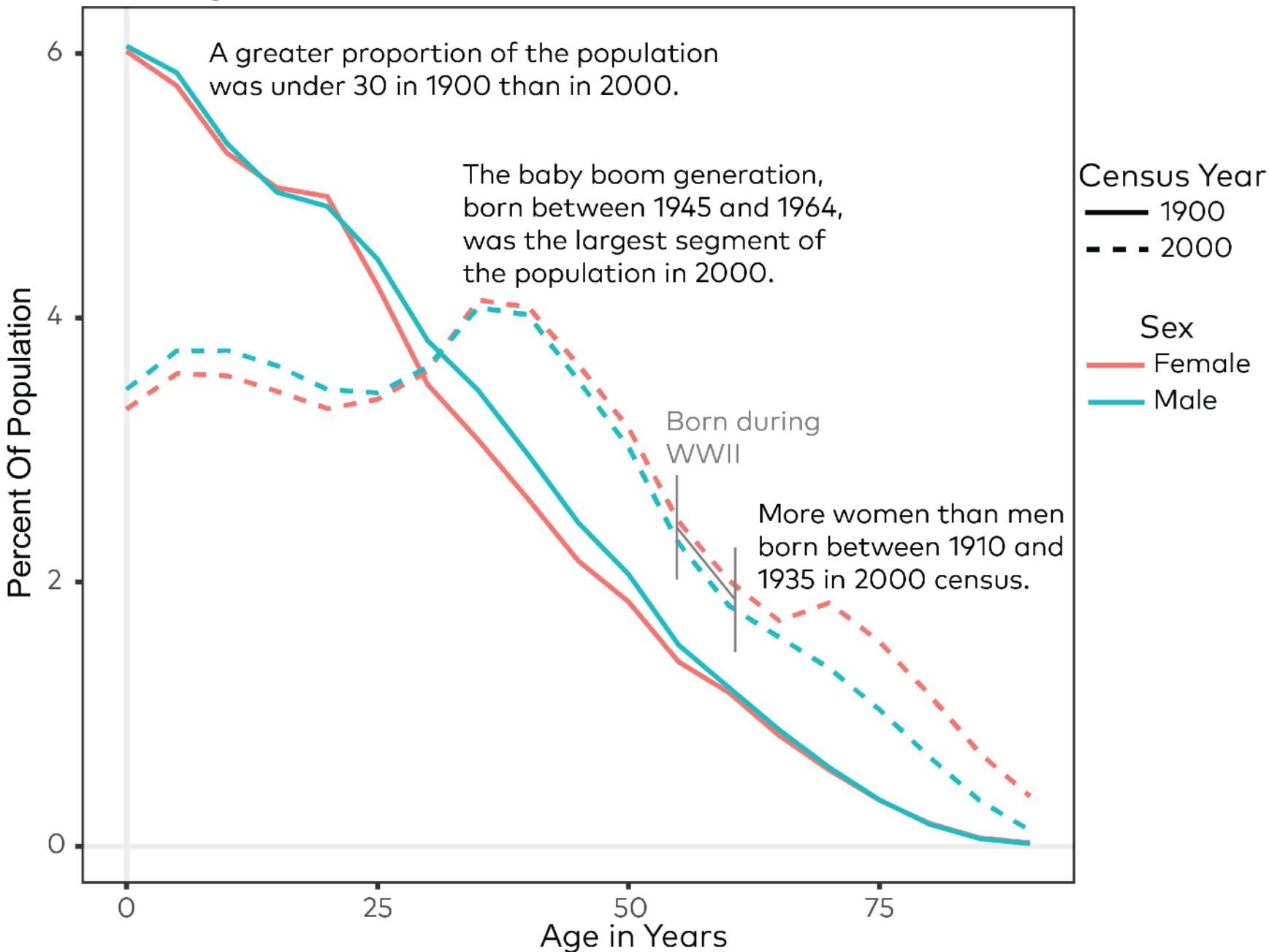
What are the differences in the male and female US population demographics, and how did they change in the 20th century?

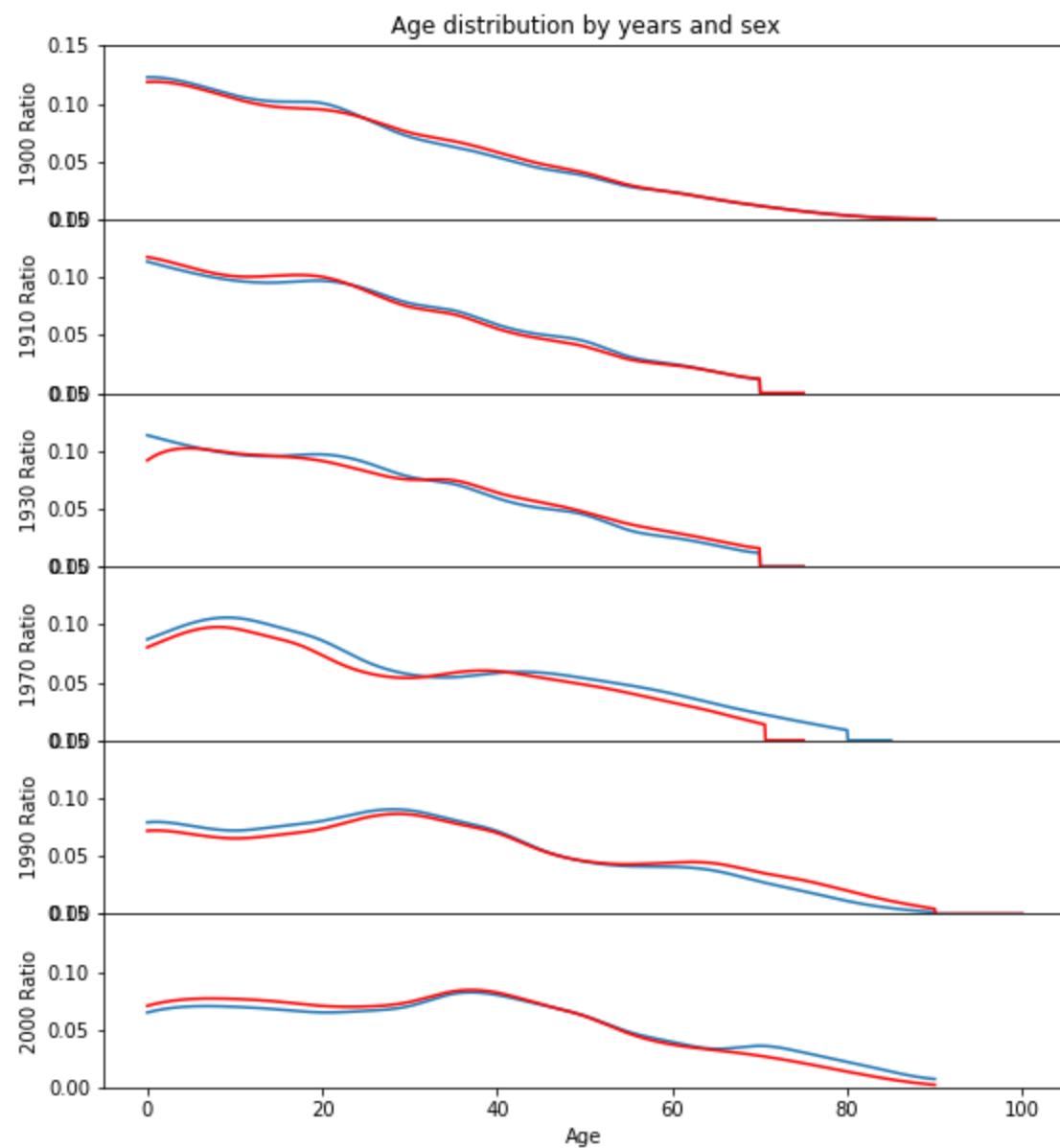


How Did Age Distribution in the United States Change Over the Last Century?

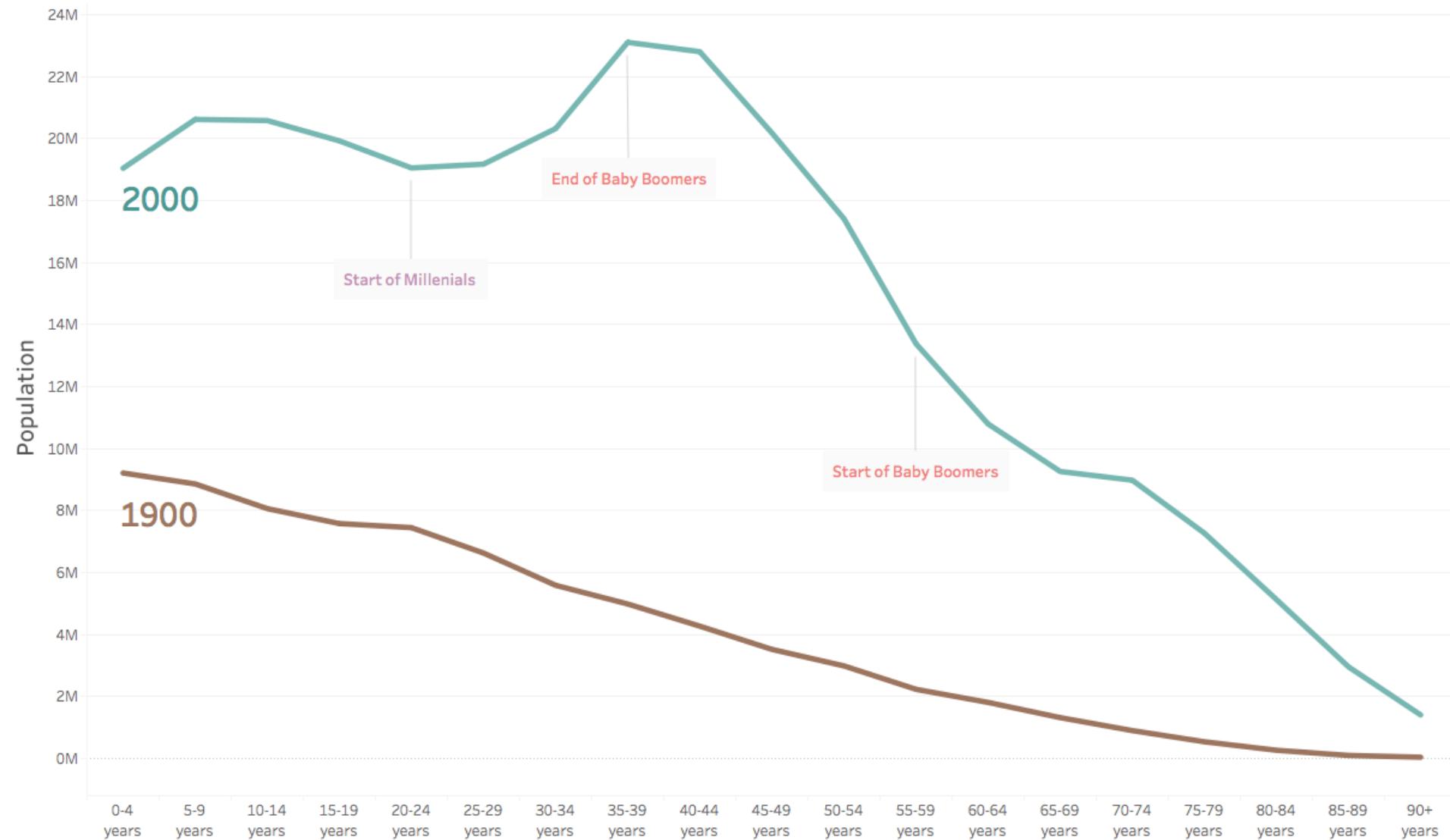


What Were the Proportions of the Population by Age and Sex for Census Years 1900 and 2000?

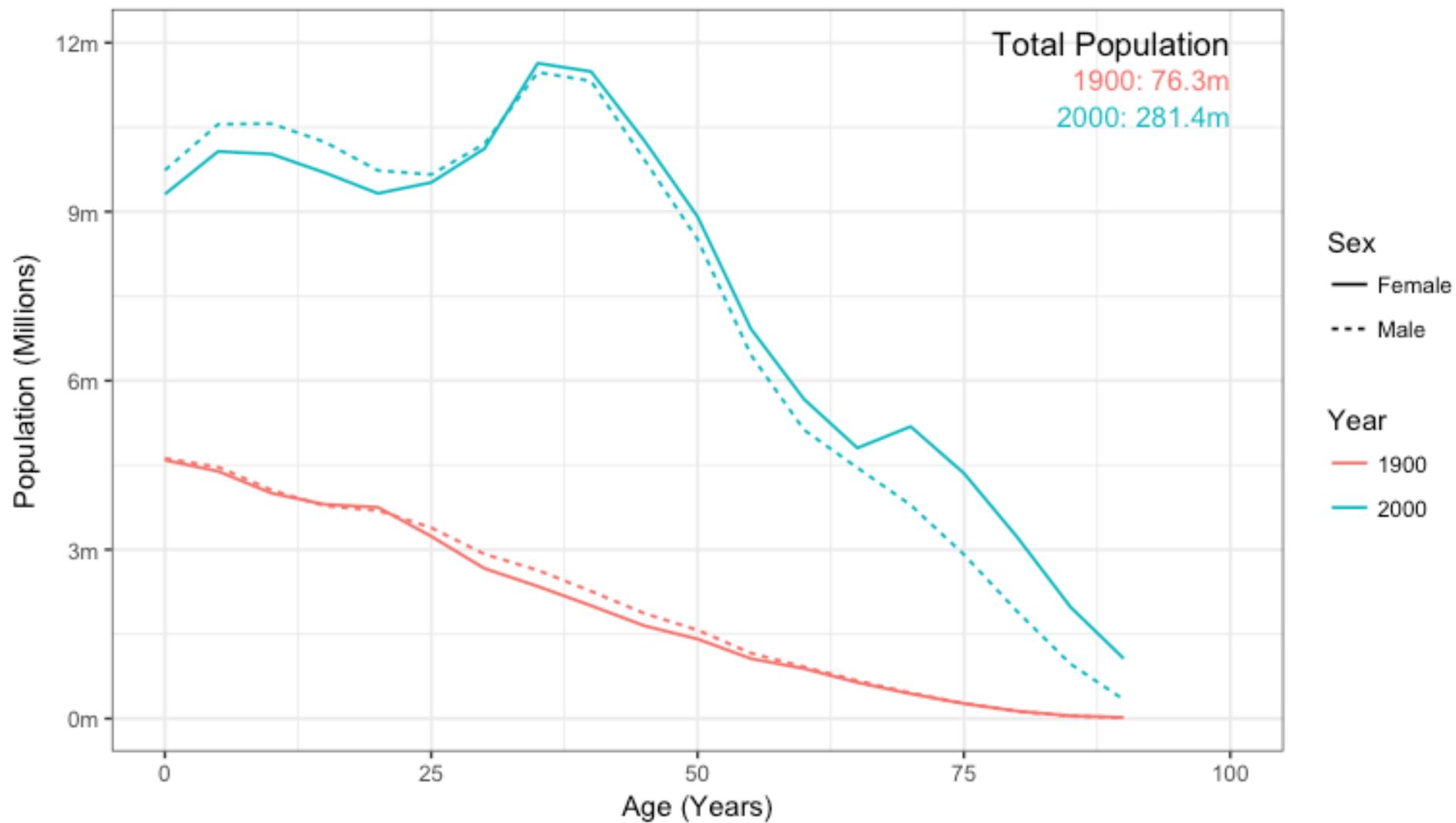




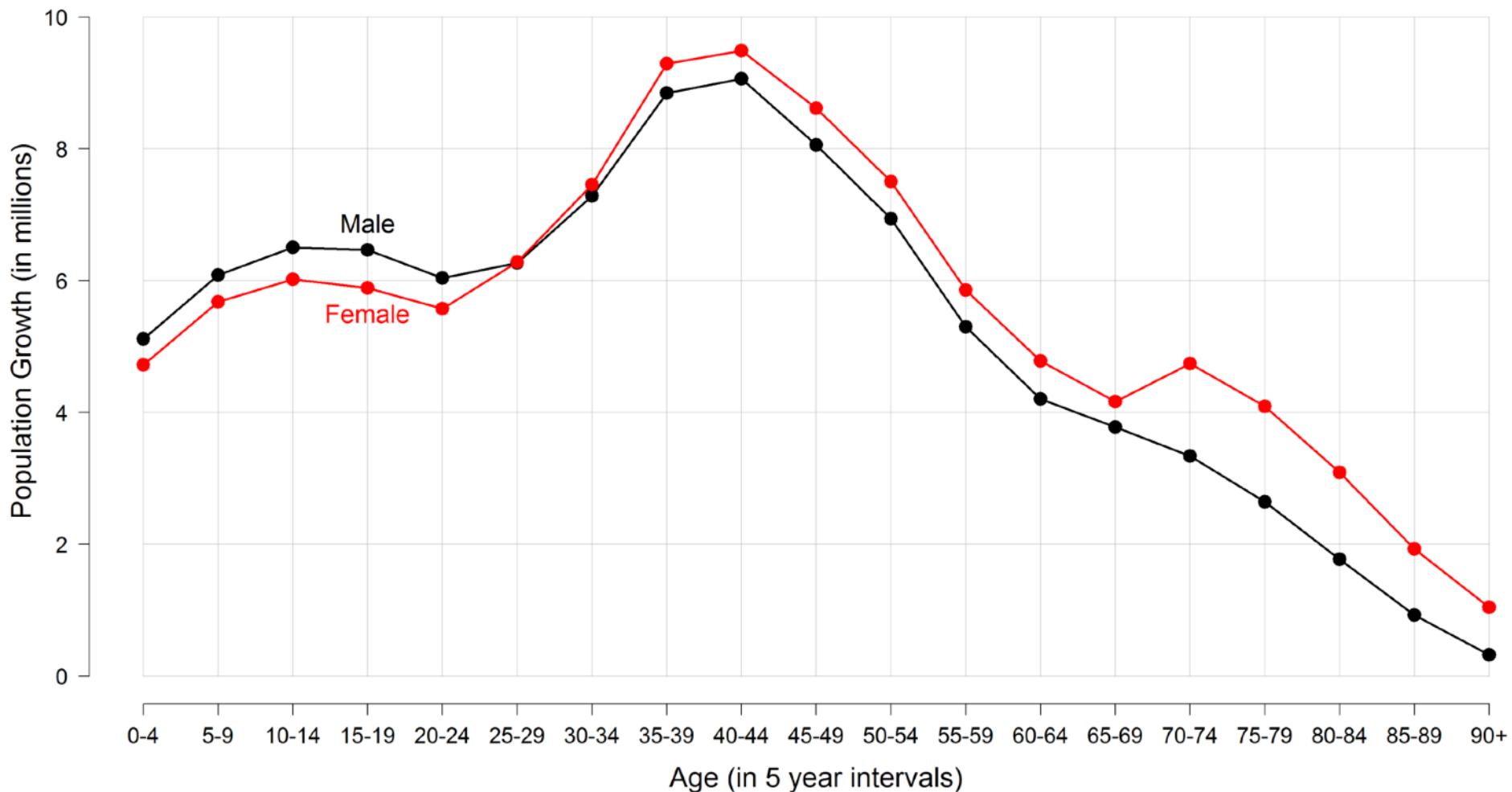
How has the age distribution in the U.S. changed since 1900?



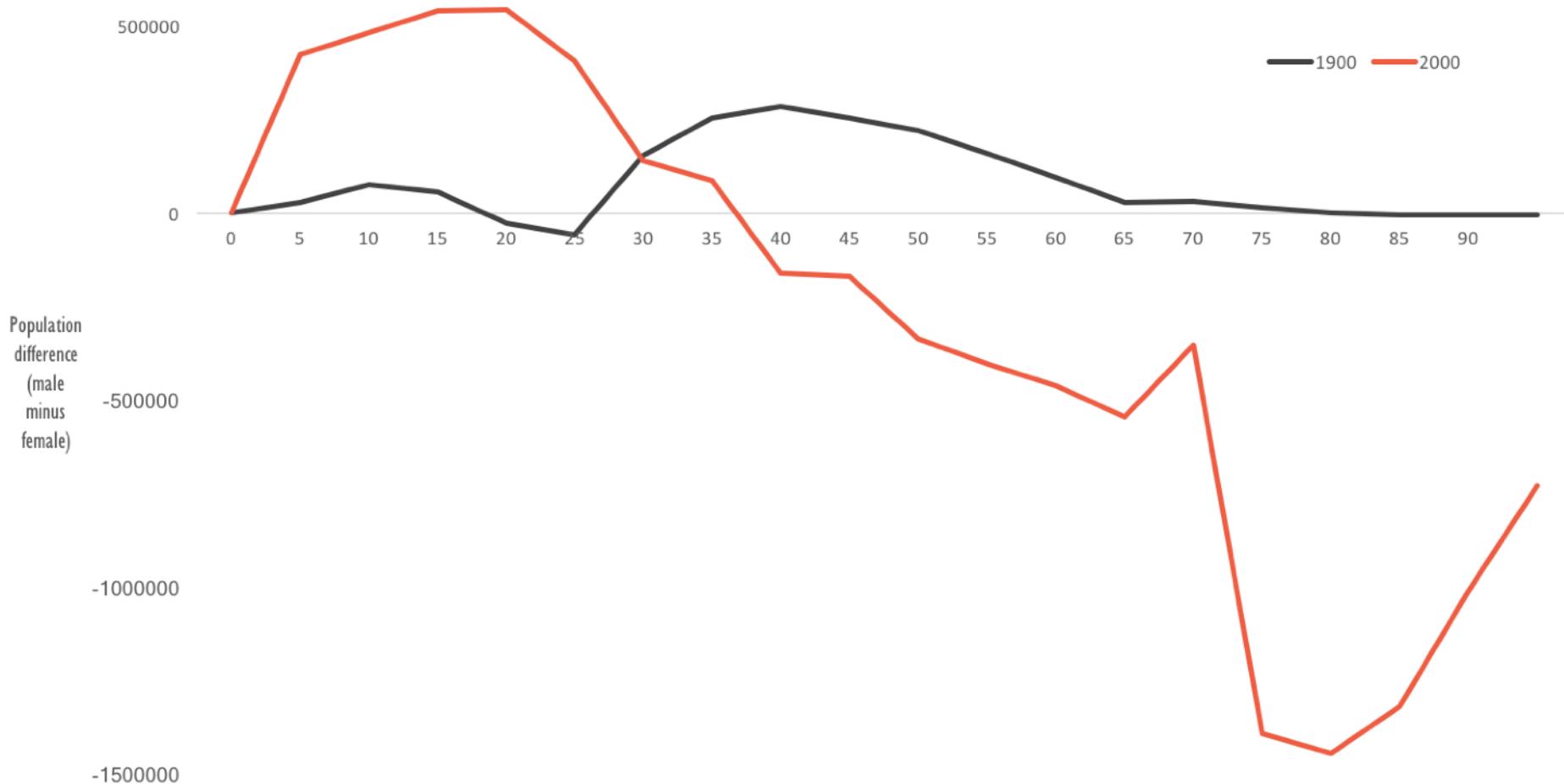
United States Population Census Bureau Data for 1900 and 2000



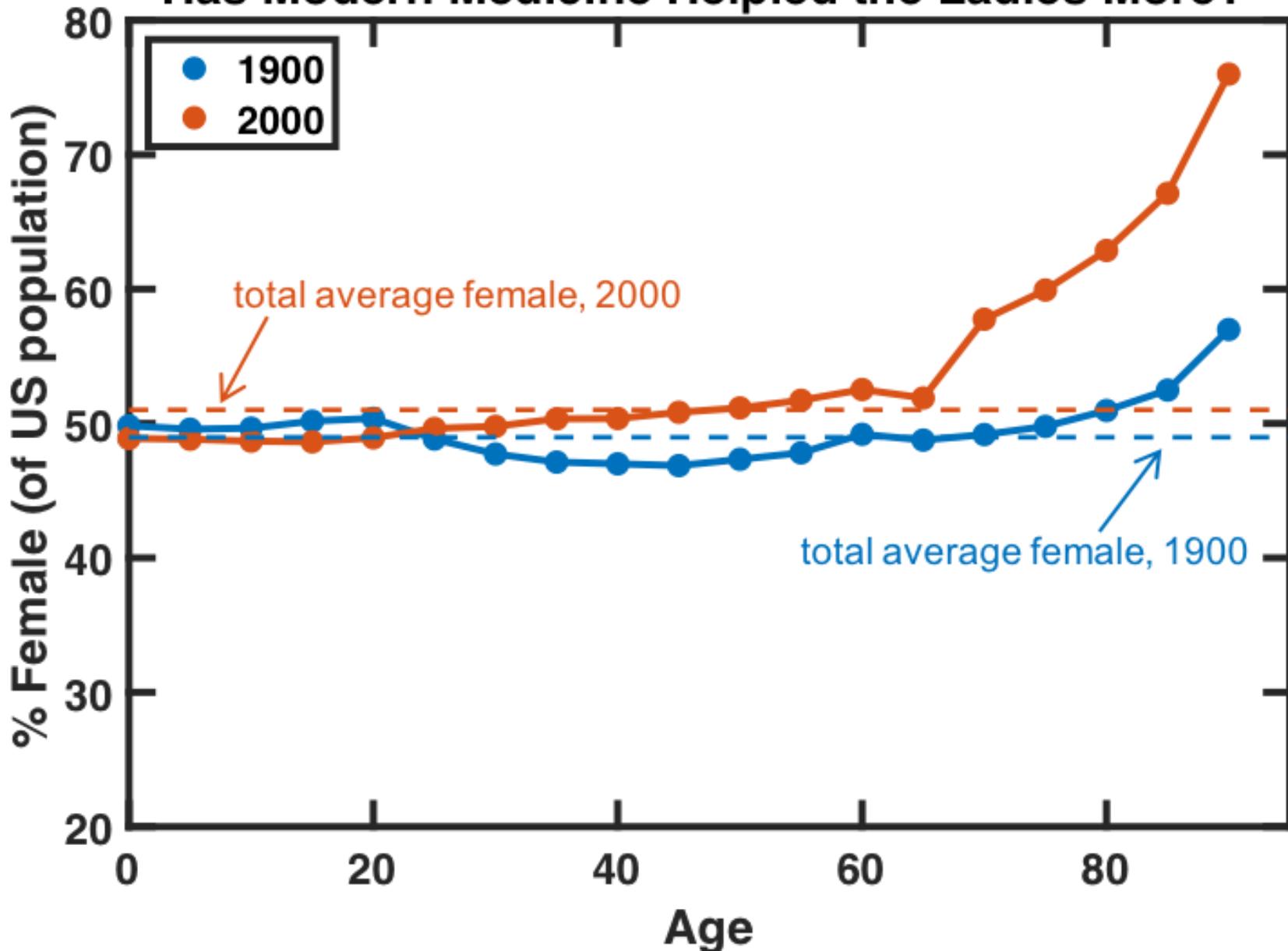
How does the U.S. population growth between 1900 and 2000 vary by age and sex?



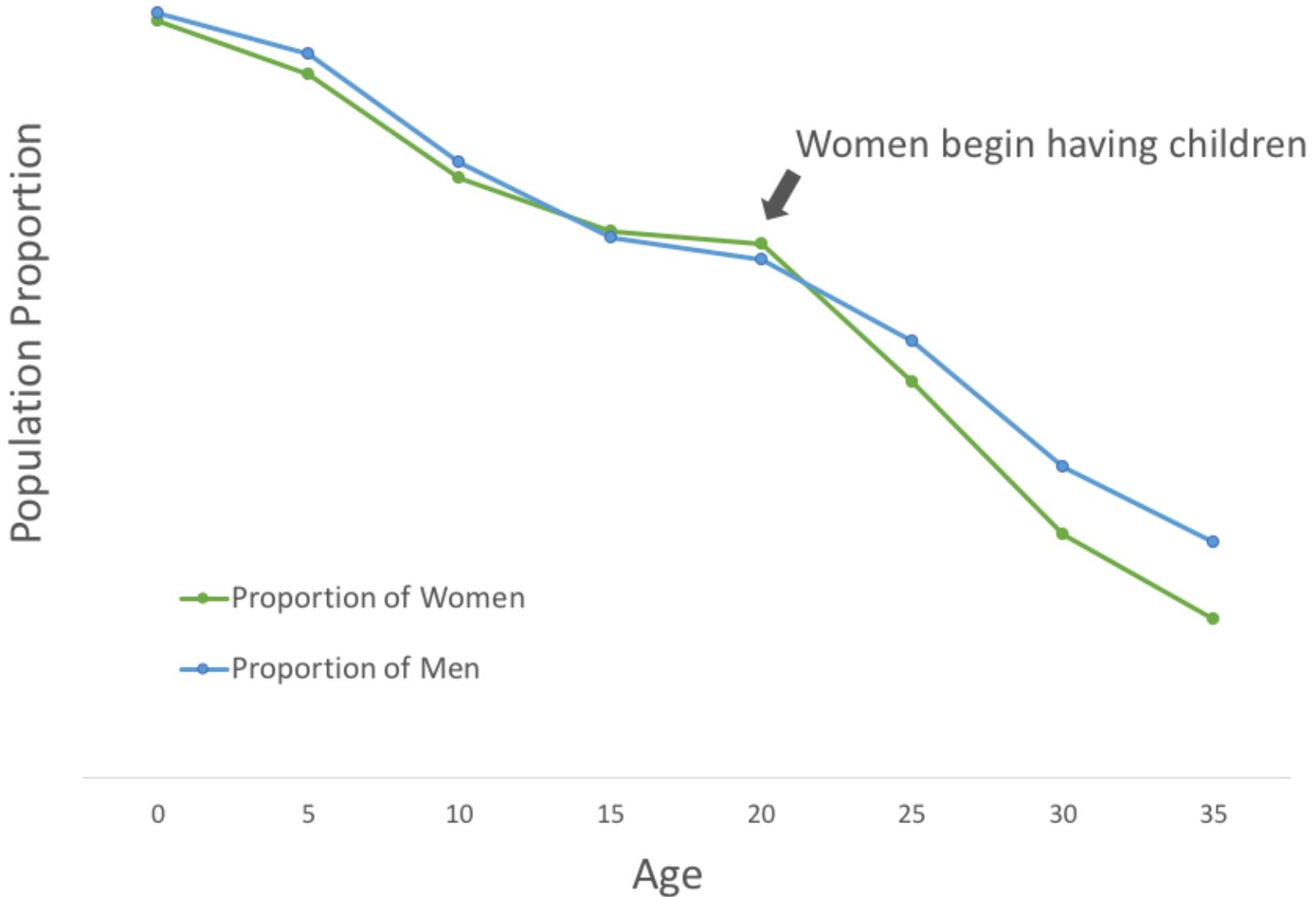
How do genders differ by age group in 1900 and 2000?



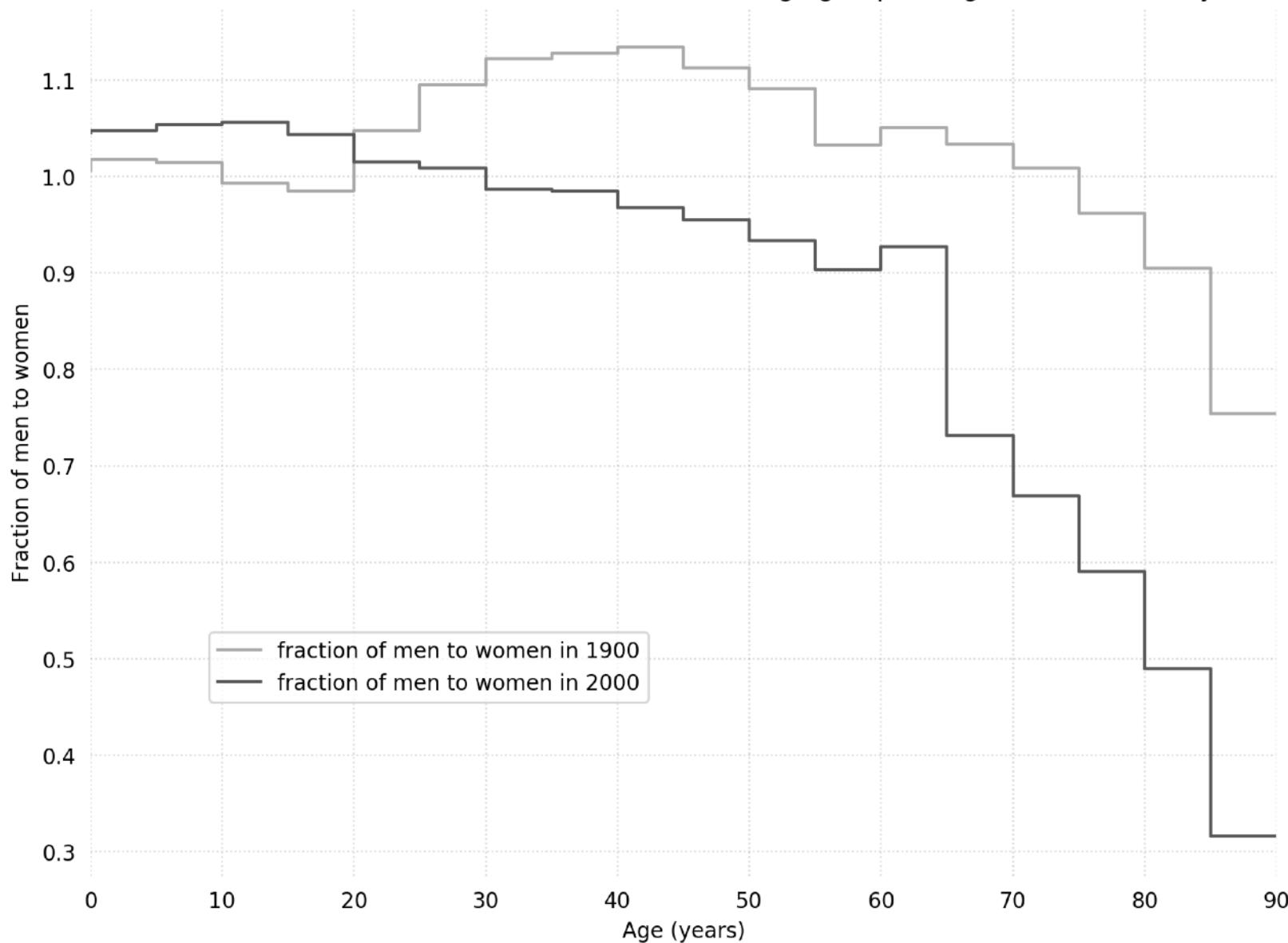
Has Modern Medicine Helped the Ladies More?



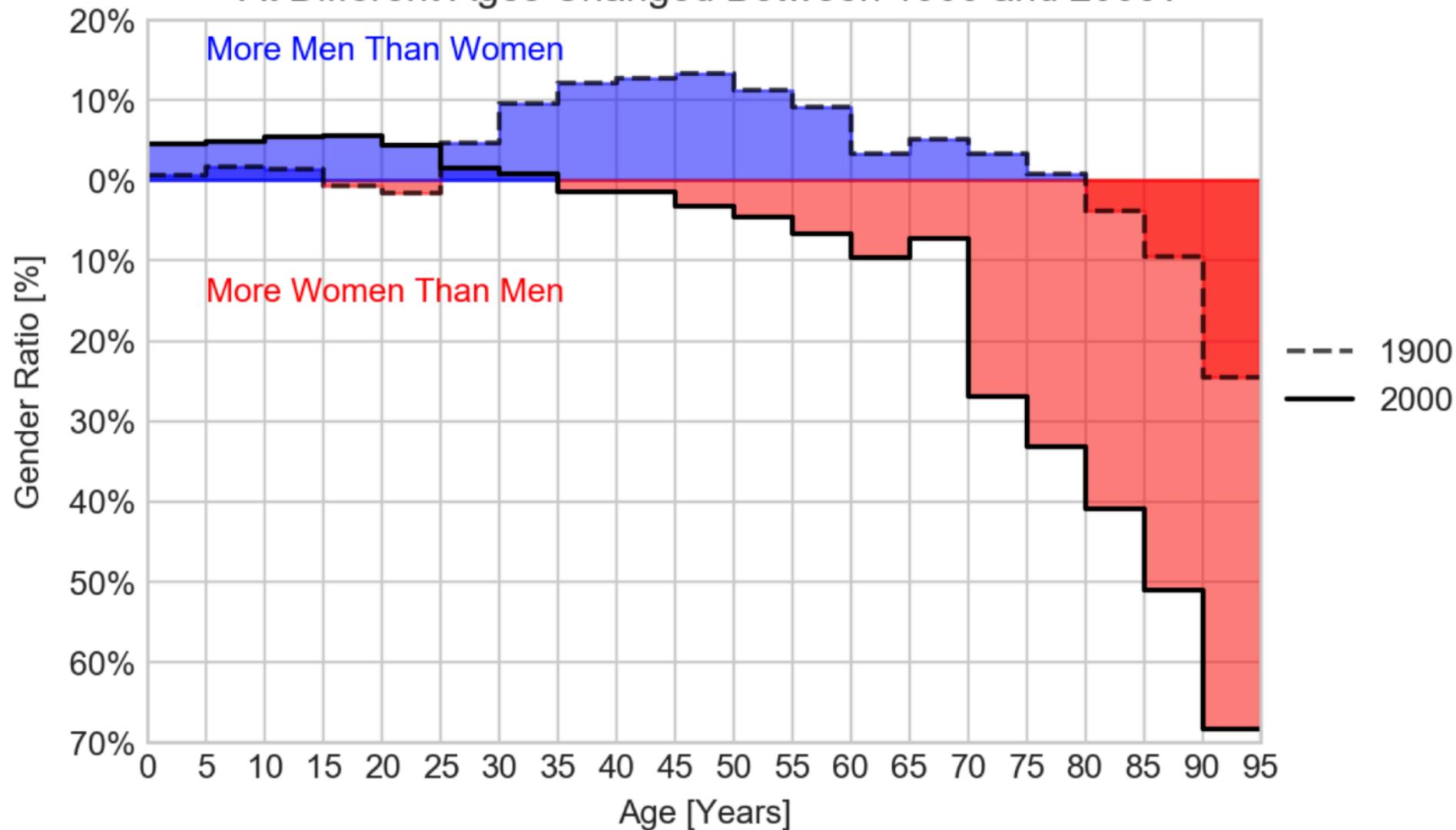
The Danger of Having Children in the Late 1800's



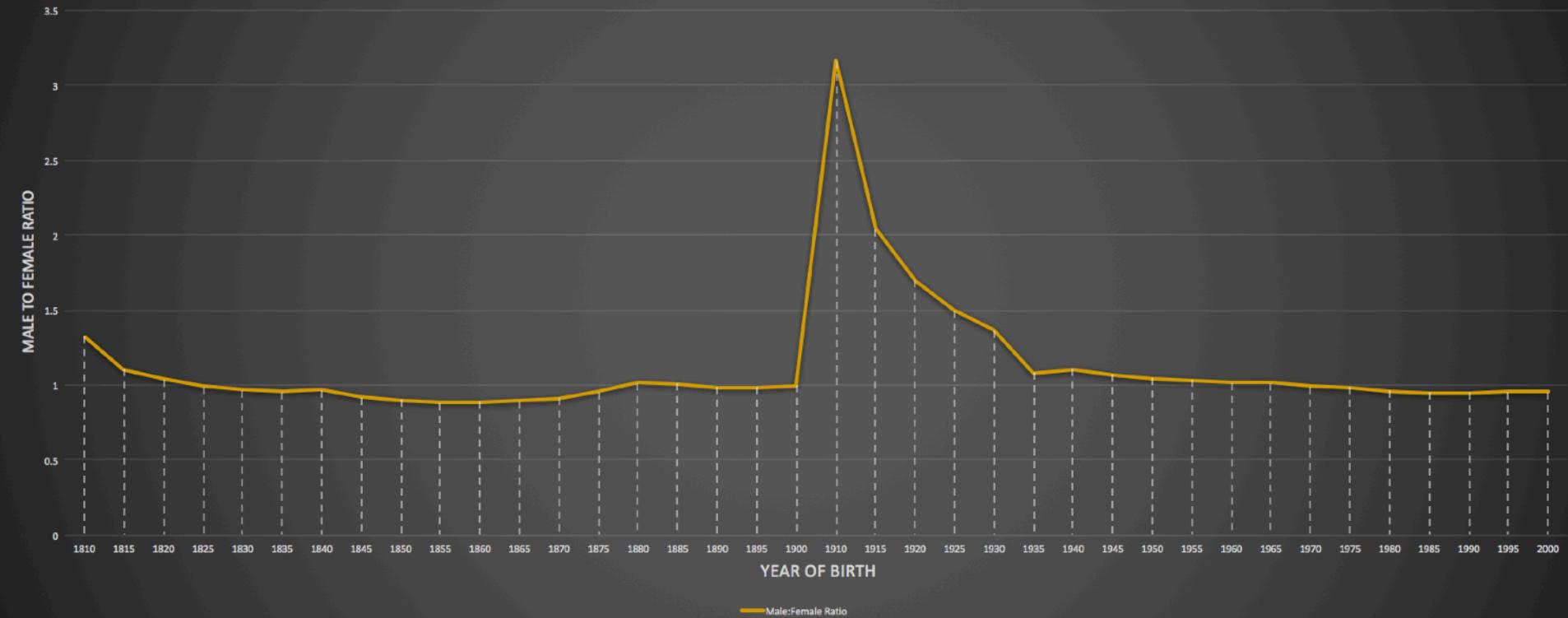
How does the fraction of men to women at each age group change across a century?



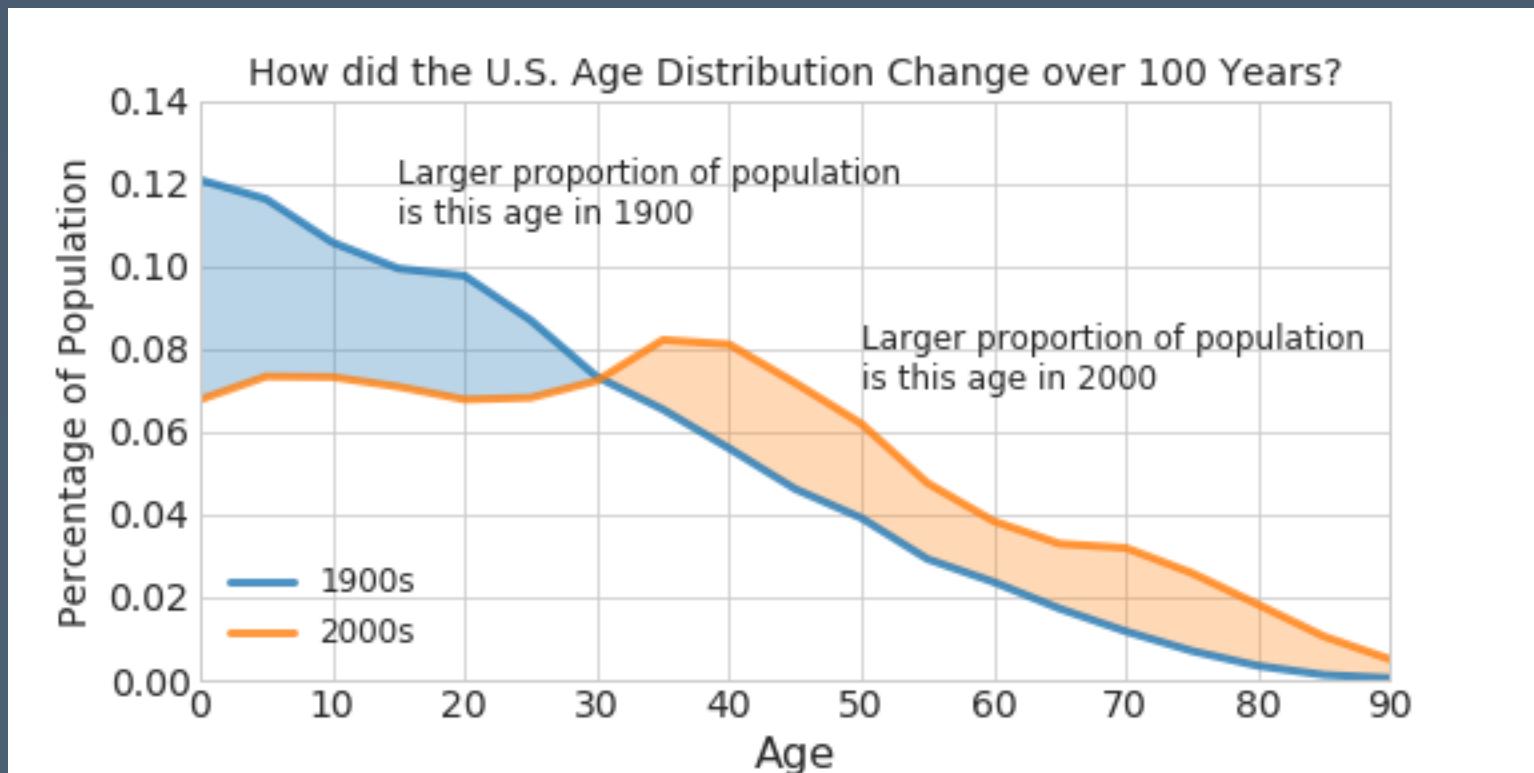
How Has The Number of Men Compared to Women At Different Ages Changed Between 1900 and 2000?



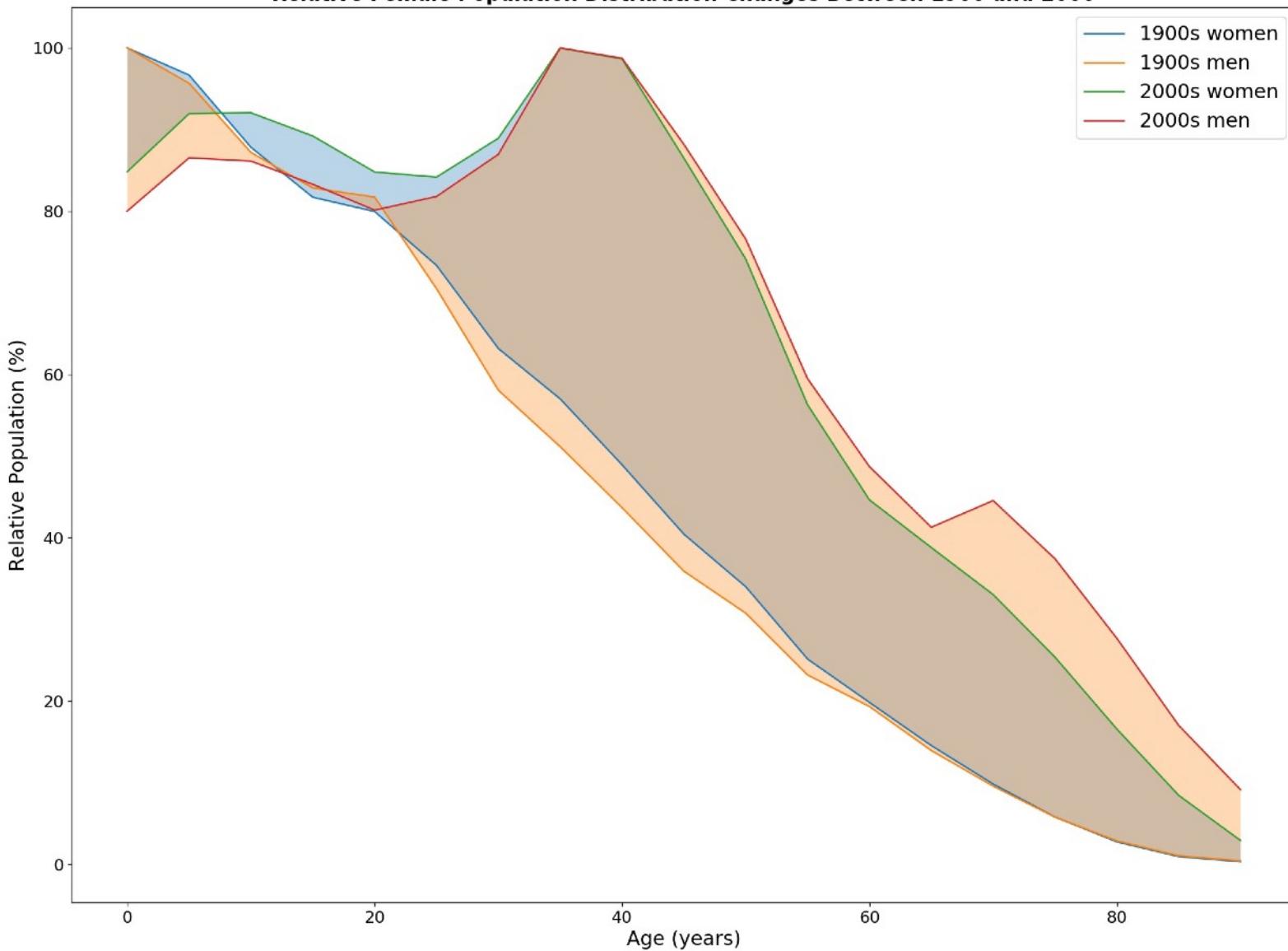
How has the Male to Female Ratio changed over the years from 1810 to 2000 ?



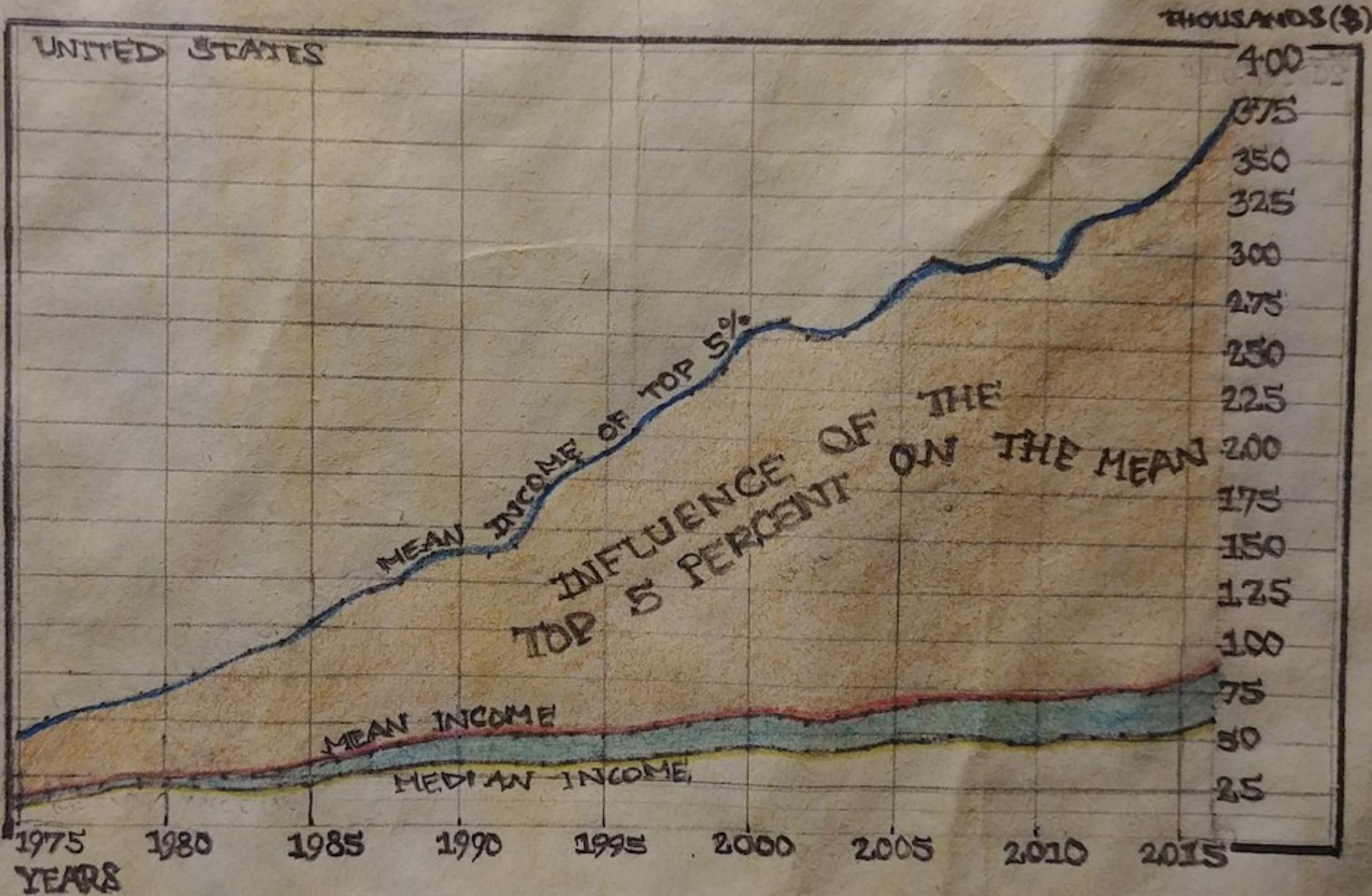
Ode to Playfair



Relative Female Population Distribution Changes Between 1900 and 2000

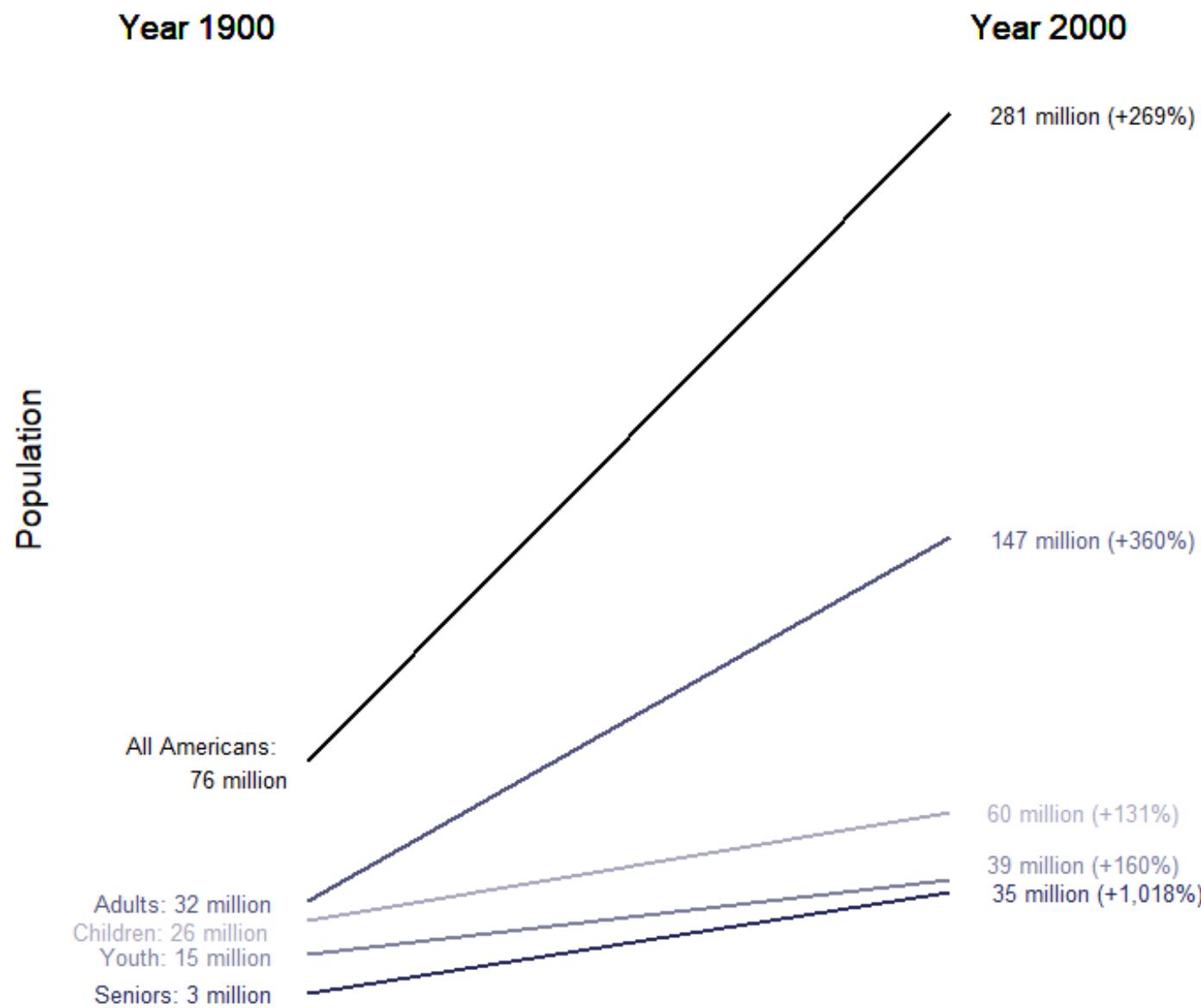


A TALE OF TWO AVERAGE INCOMES

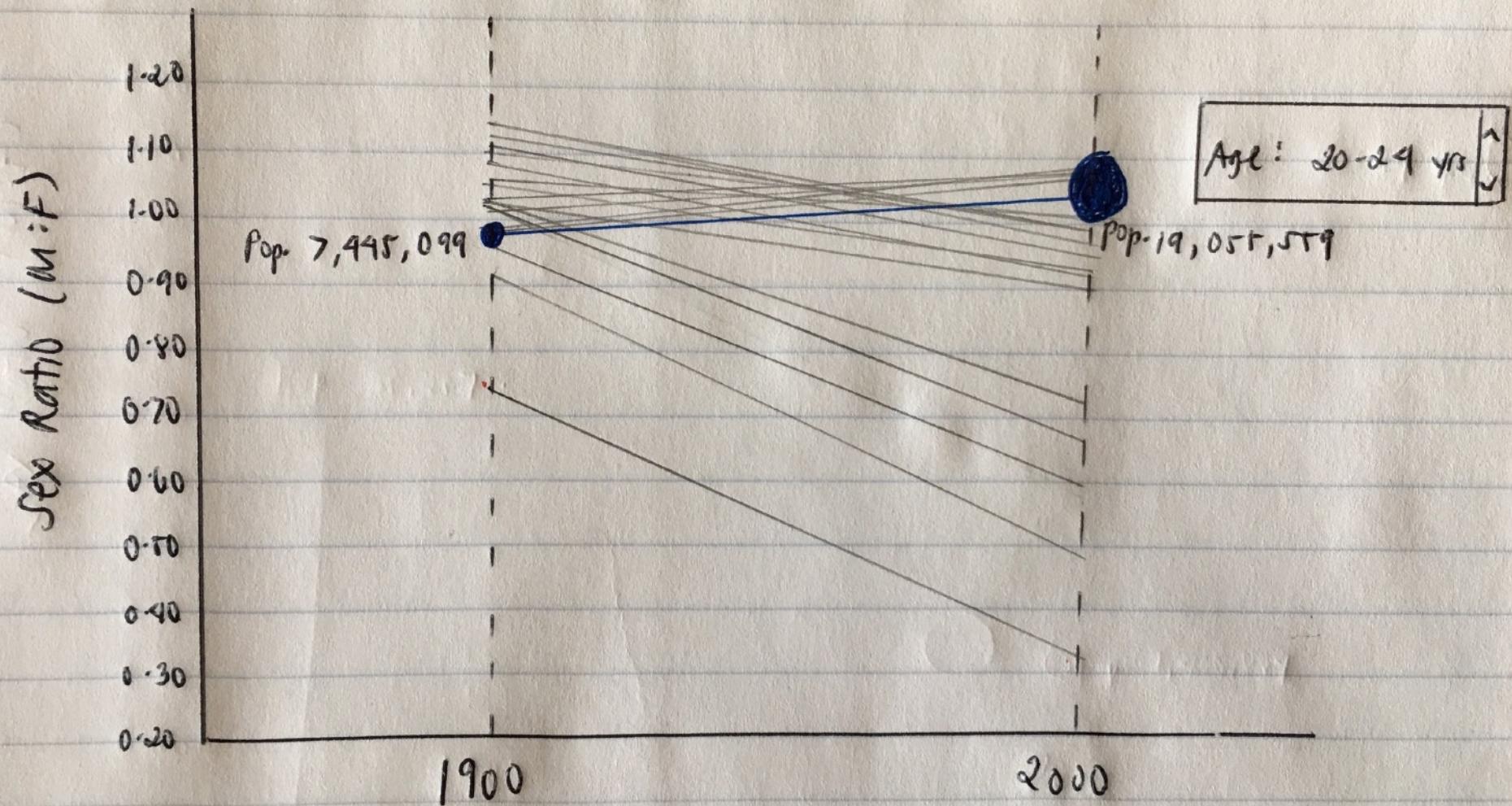


Slope Graph

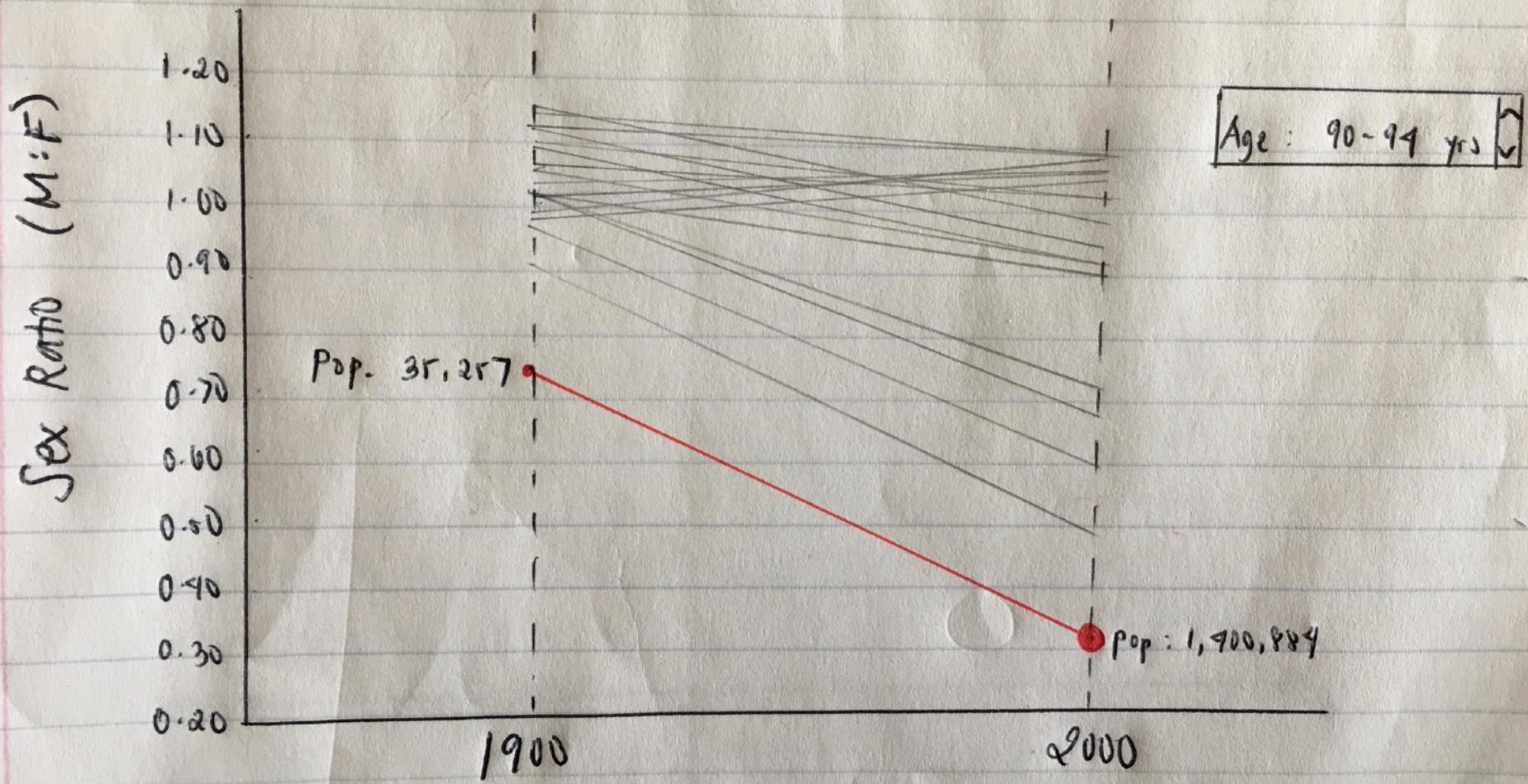
How did the population growth of individual age groups contribute to the overall population growth in the United States between 1900 and 2000?



Sex Ratio Change Between 1900 and 2000

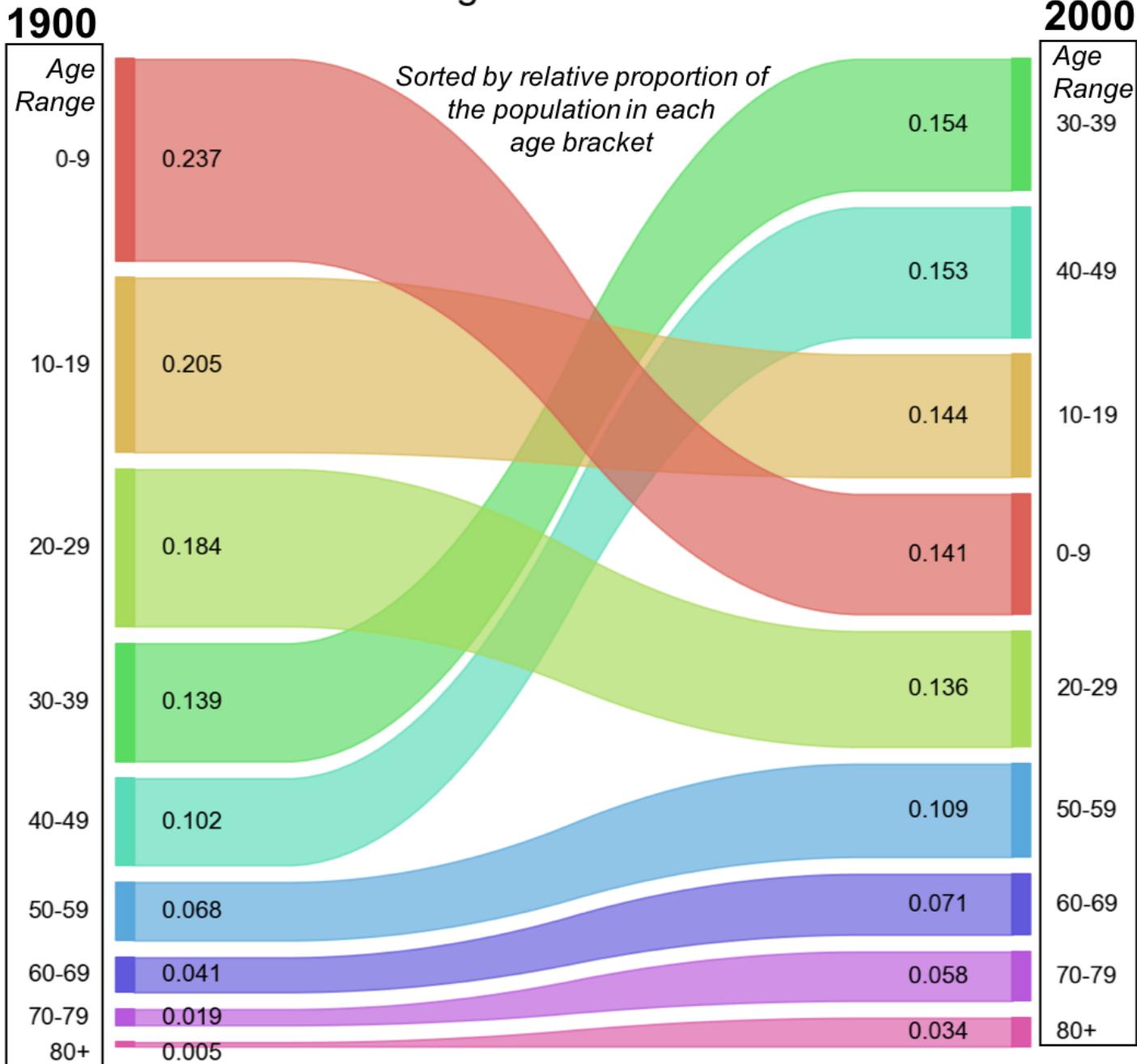


Sex Ratio Change Between 1900 and 2000



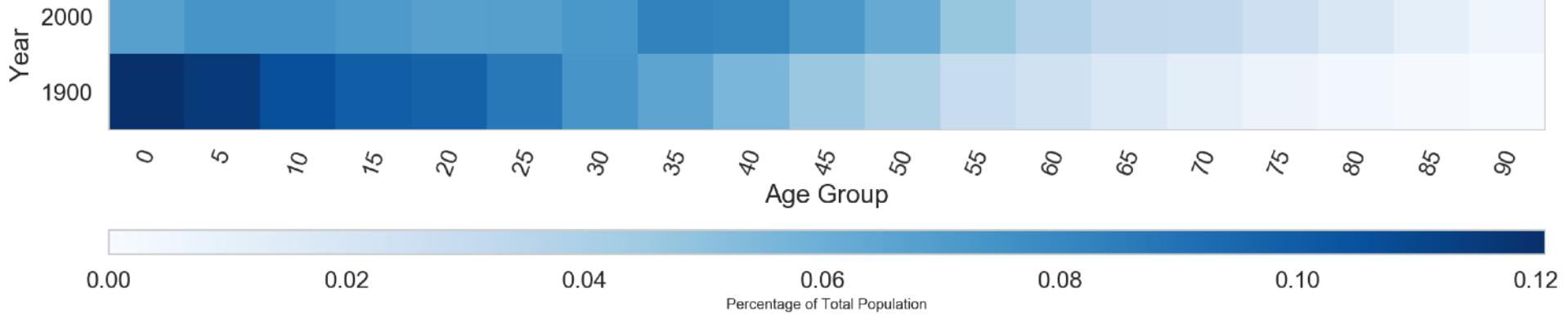
Sankey Diagram

Did the relative proportion of age demographics in the US change between 1900 and 2000?



Heatmaps

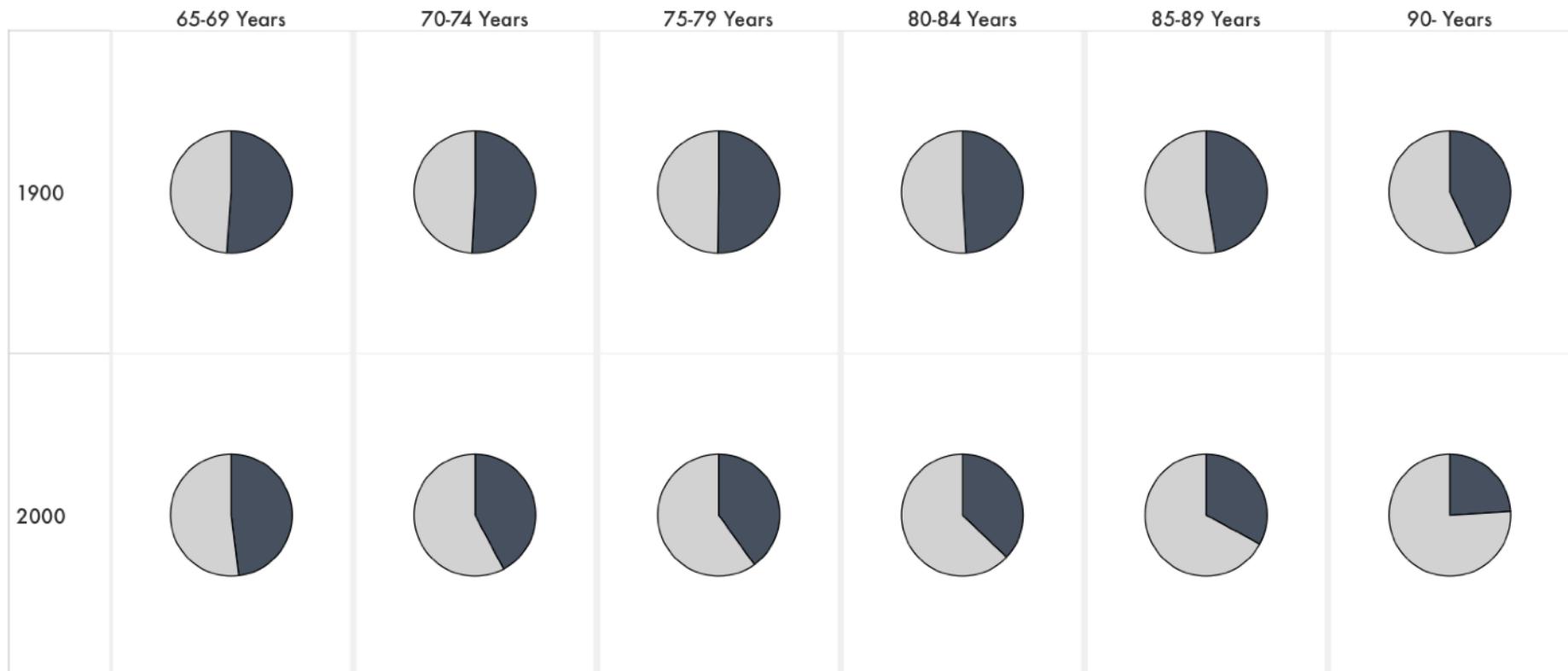
Has the age group containing the largest percentage of Americans shifted?



Pie Charts

How do different age groups of seniors (65+) differ by gender composition for 1900 and 2000?

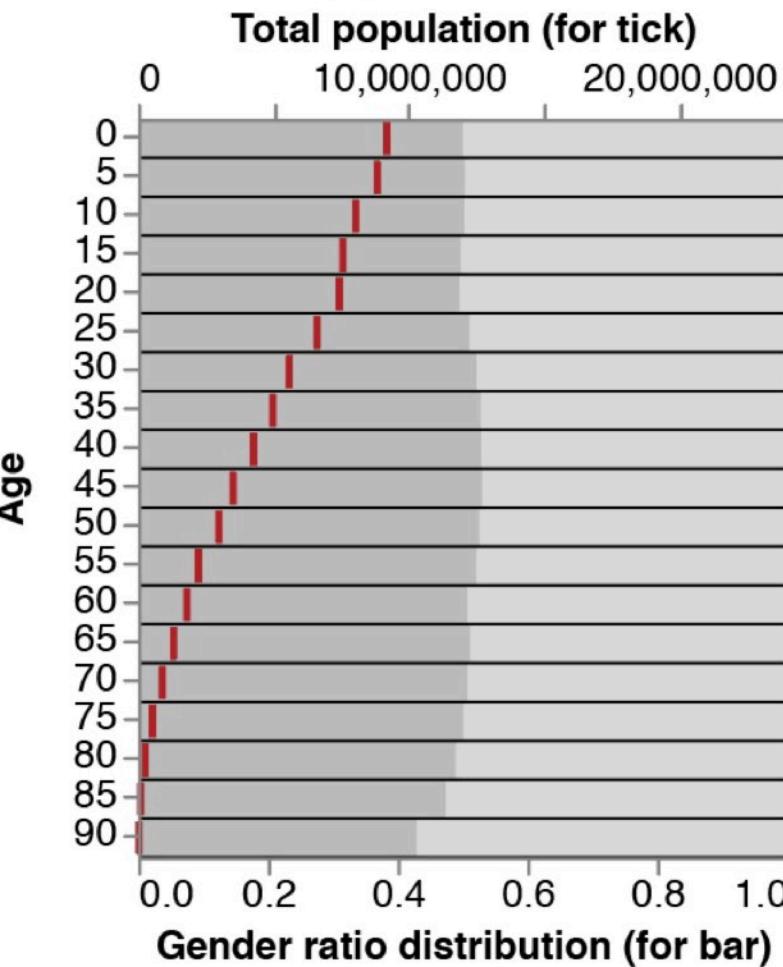
Legend
□ % Female
■ % Male



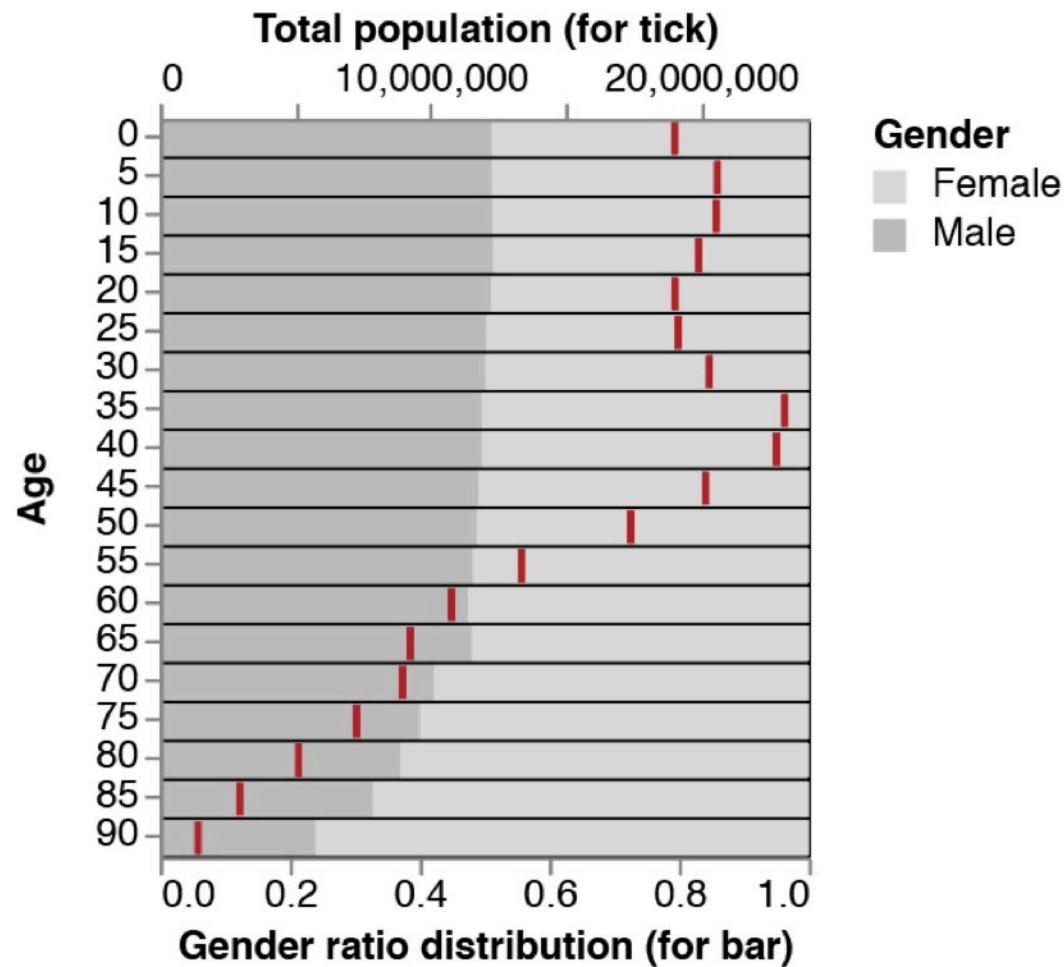
Multi-View or Dual Axis

What is population distribution in 1900 and 2000 with respect to age?

(a) Year: 1900



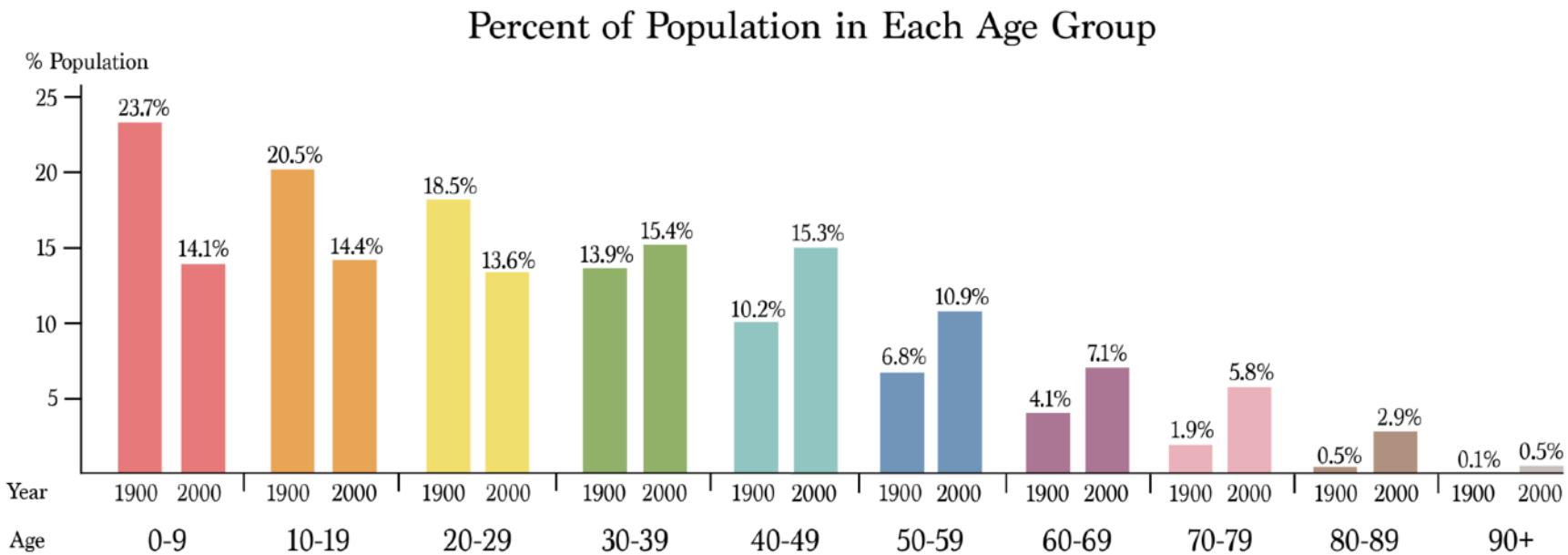
(b) Year: 2000



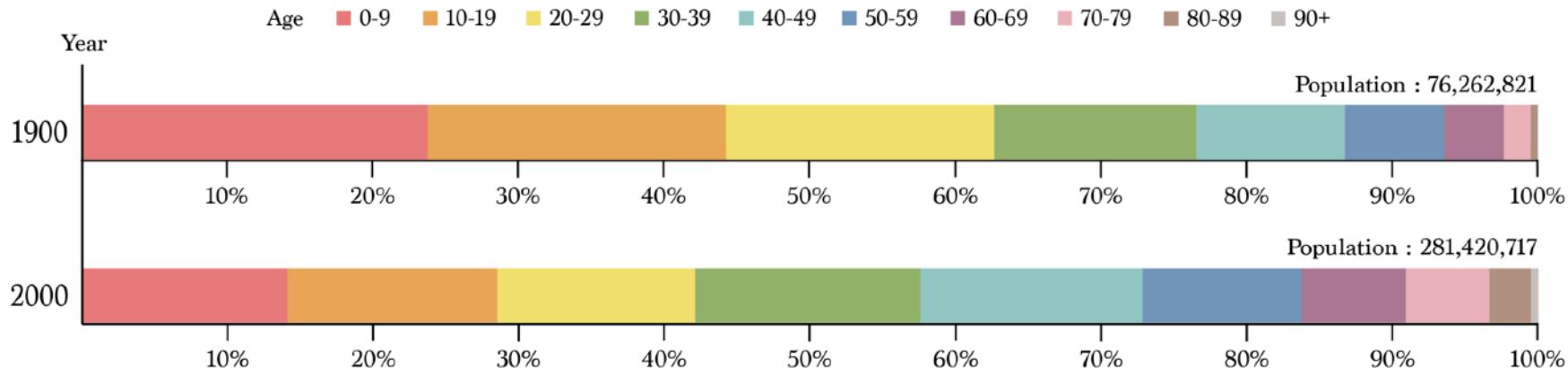
Gender

- Female
- Male

How has the Distribution of Age in the U.S. Changed from 1900 to 2000?



Age Composition of Total Population



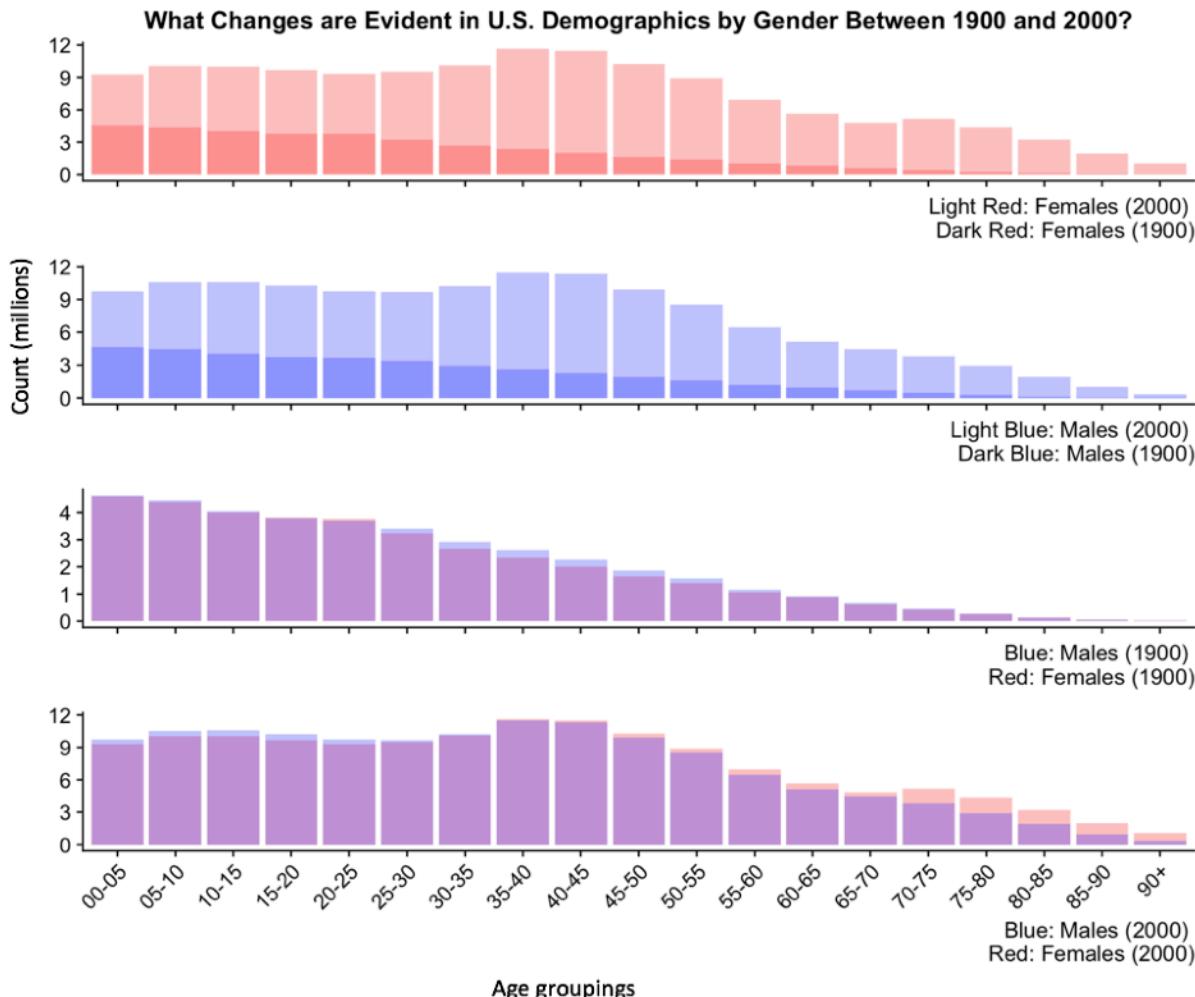
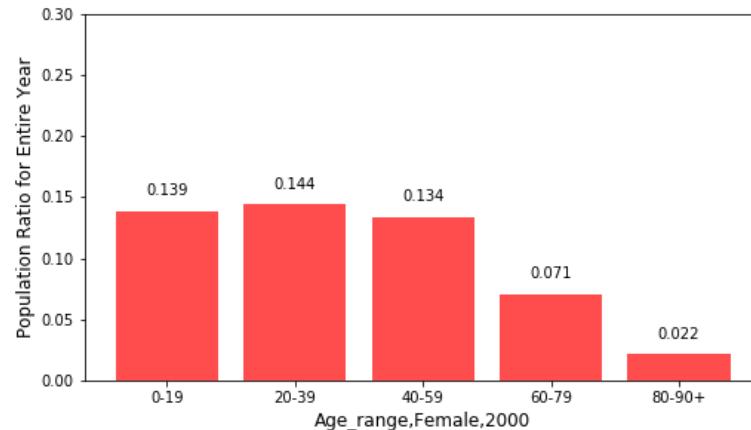
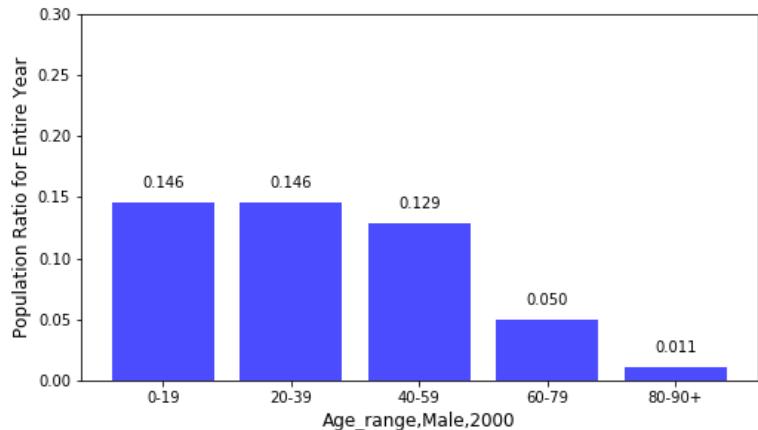
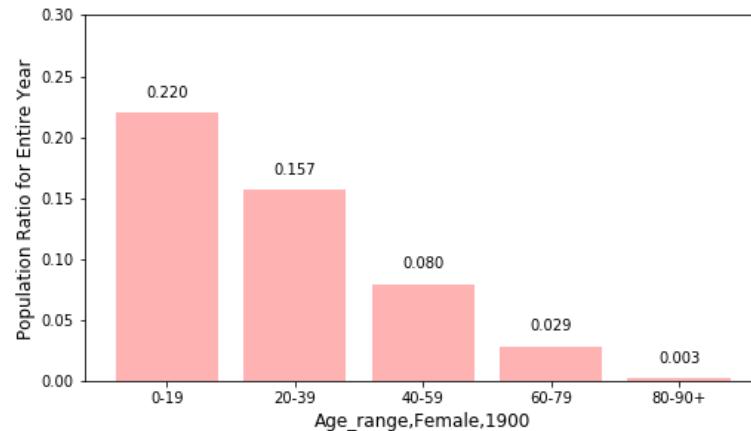
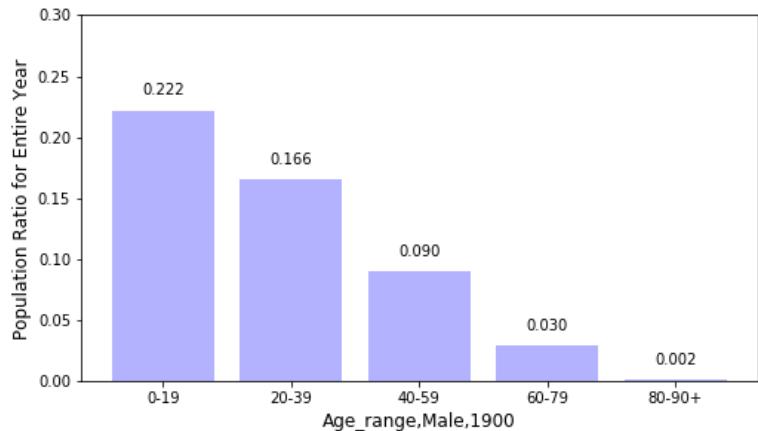


Figure 1: U.S. Demographics by Gender in 1900 and 2000 (Source: U.S. Census data)

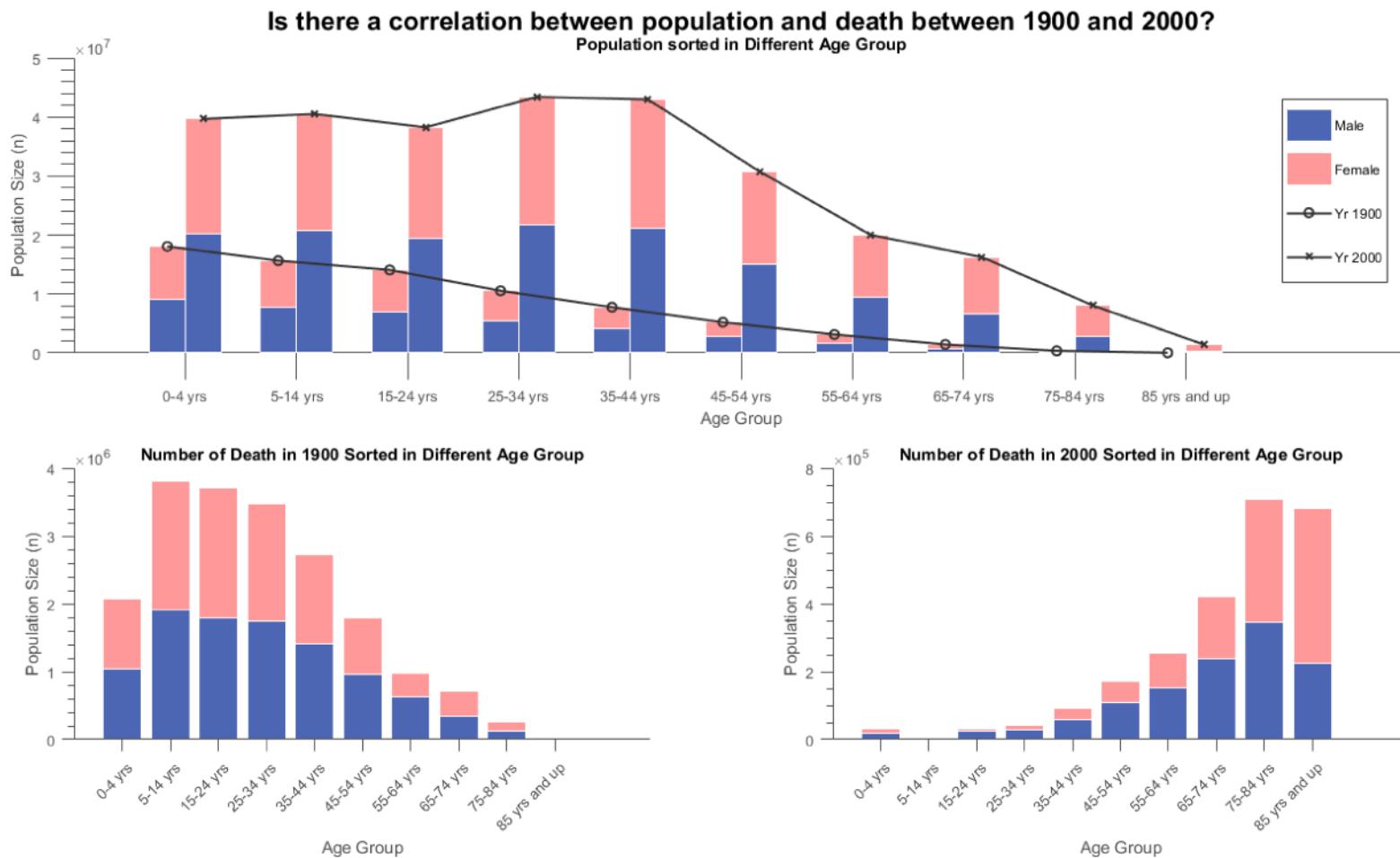
NOTE: Overlapping histograms are used to compare data sets in each plot. Each bin includes values less than the final age listed for its range (e.g. a value of 10 would be included in the 10-15 bin).

Plot descriptions: (top) Displays the age distribution for females in 1900 and 2000, showing the increase in population across all age groups. (second from top) Displays the age distribution for males in 1900 and 2000, showing a similar trend to that seen in the previous plot. (third from top) Displays the age distribution for both males and females in 1900. Areas where the histograms for males and females overlap are shown in purple. This plot shows that men between the ages of 25 – 55 slightly outnumber women in the same age ranges. (bottom) Displays the age distribution for both males and females in 2000. This plot shows that women aged 45 and older outnumber men in the same age ranges.

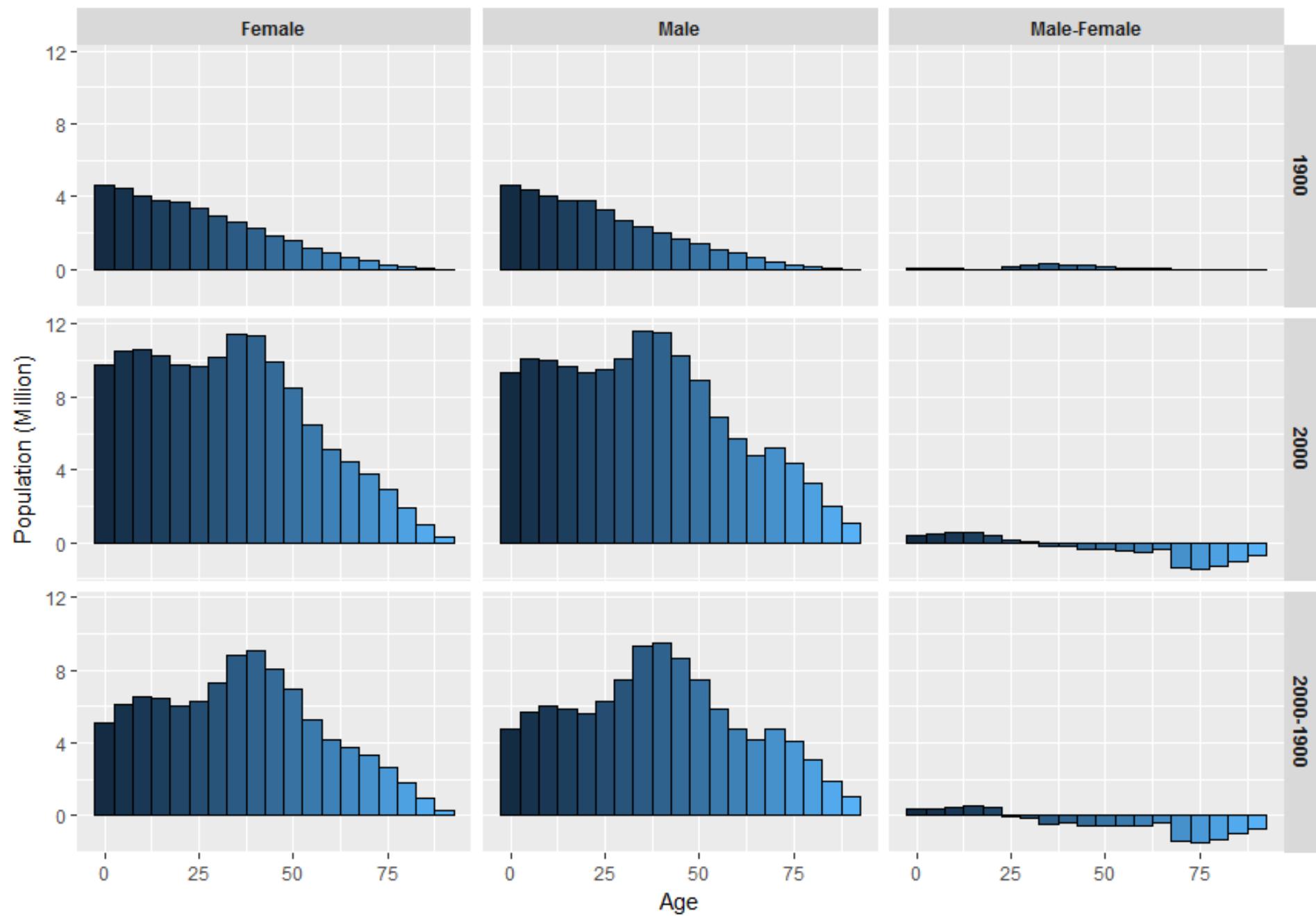
Changes of different genders from different ages between 1900 and 2000



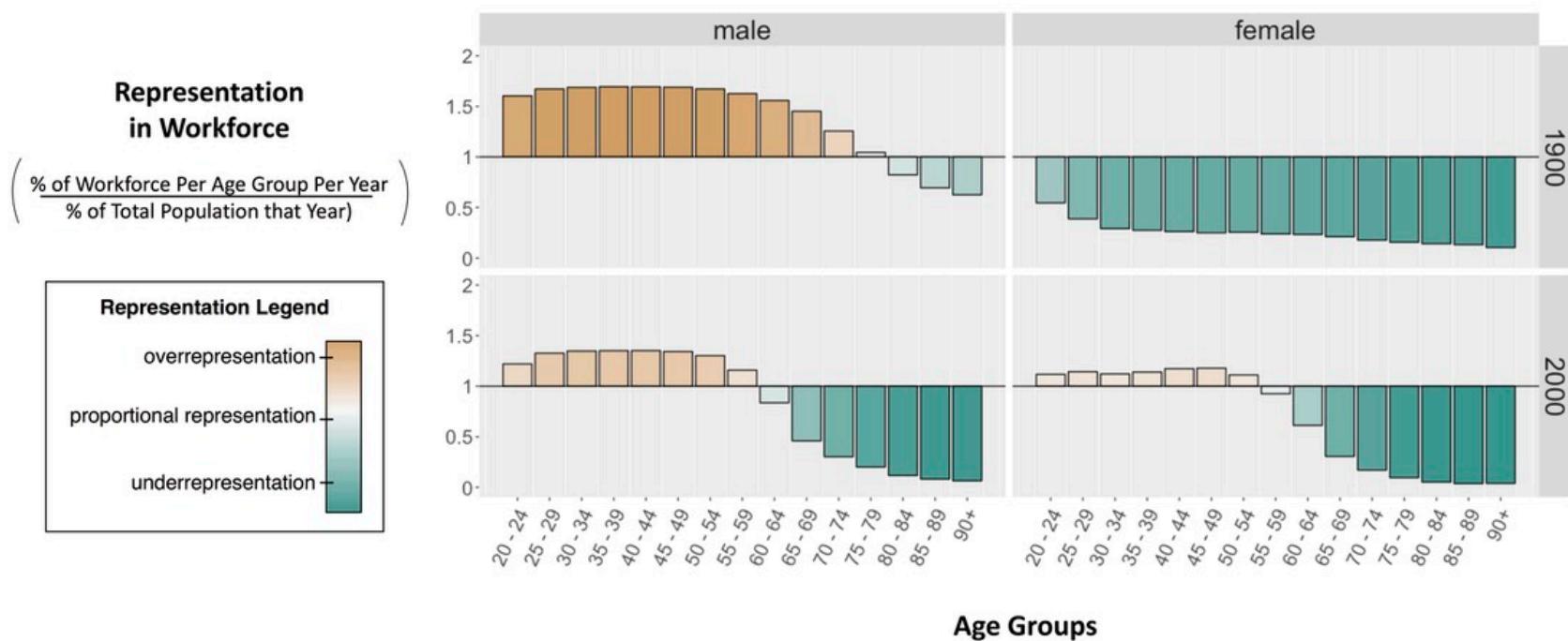
The total population of 1900 is 76262821, while of 2000 is 281420717



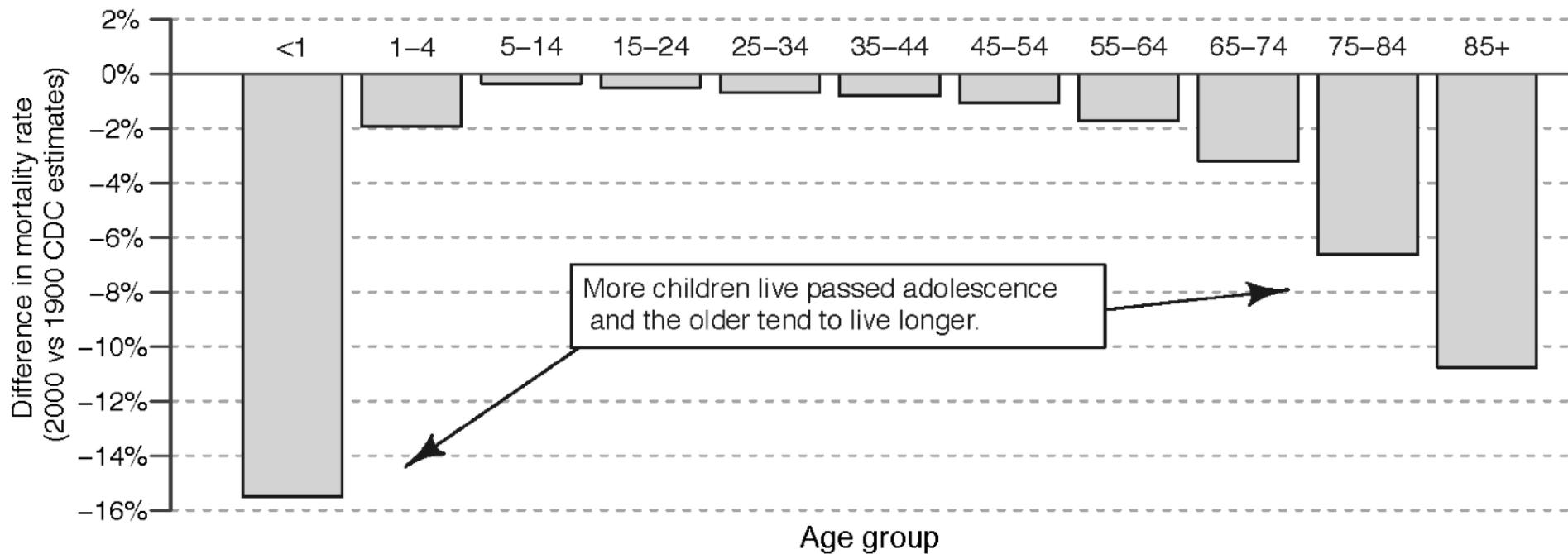
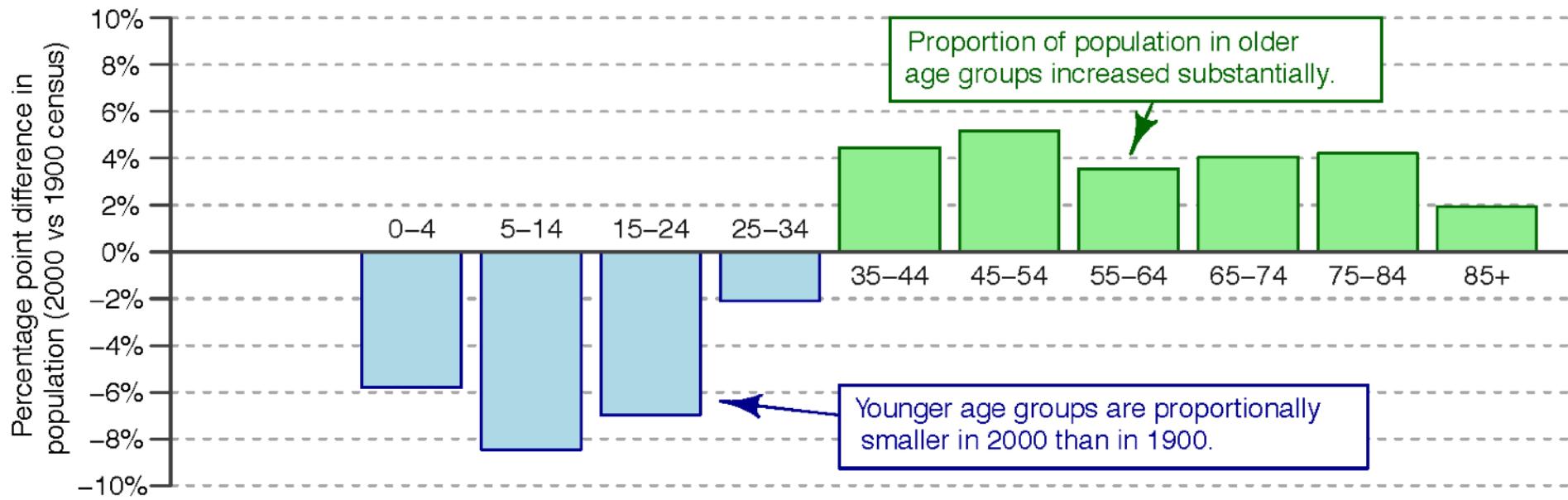
Population by Age and Gender in United States (1900 v.s. 2000)



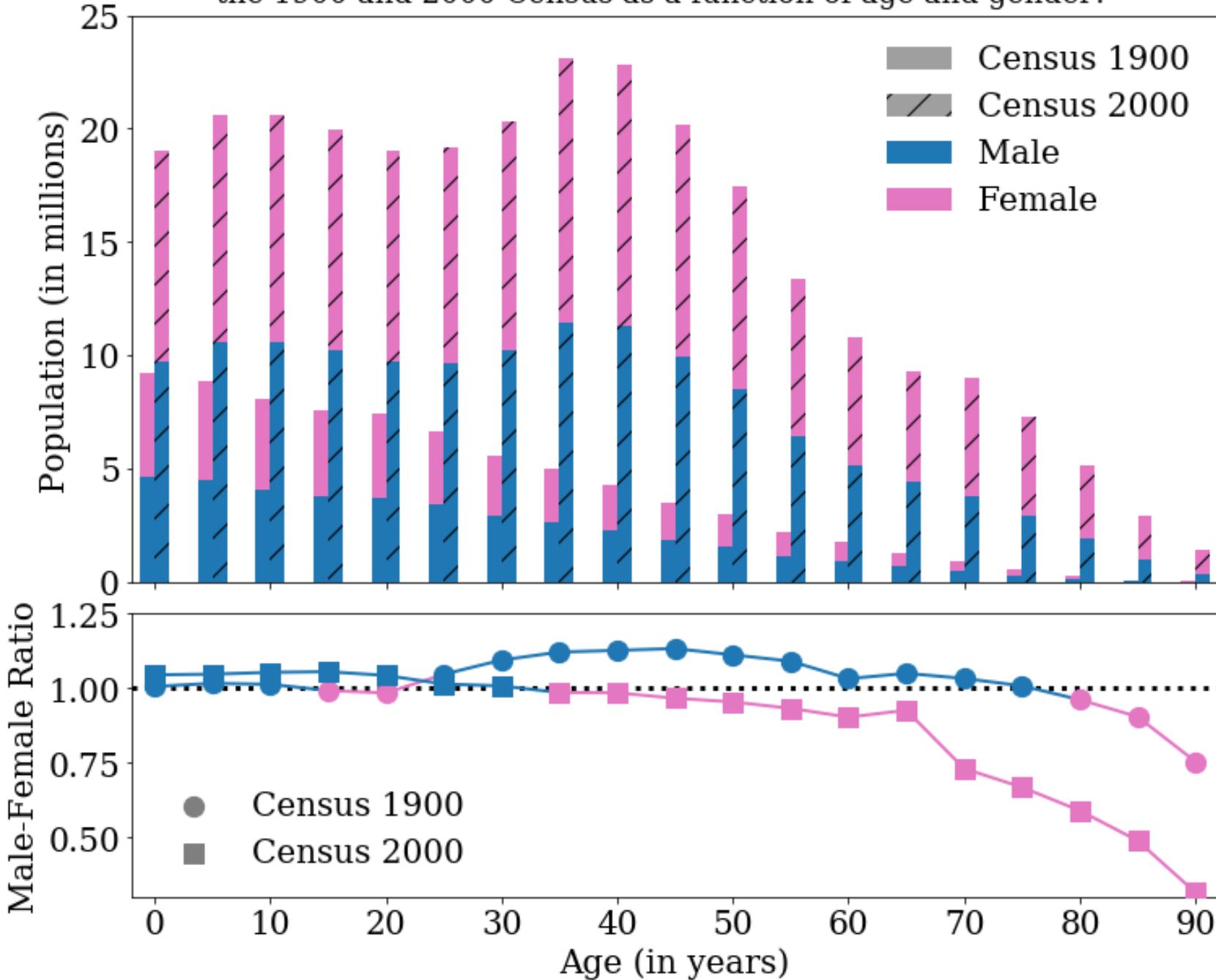
Did Gender and Age Representation in the Workforce Become More Proportional to their Respective Population Sizes Between 1900 and 2000?



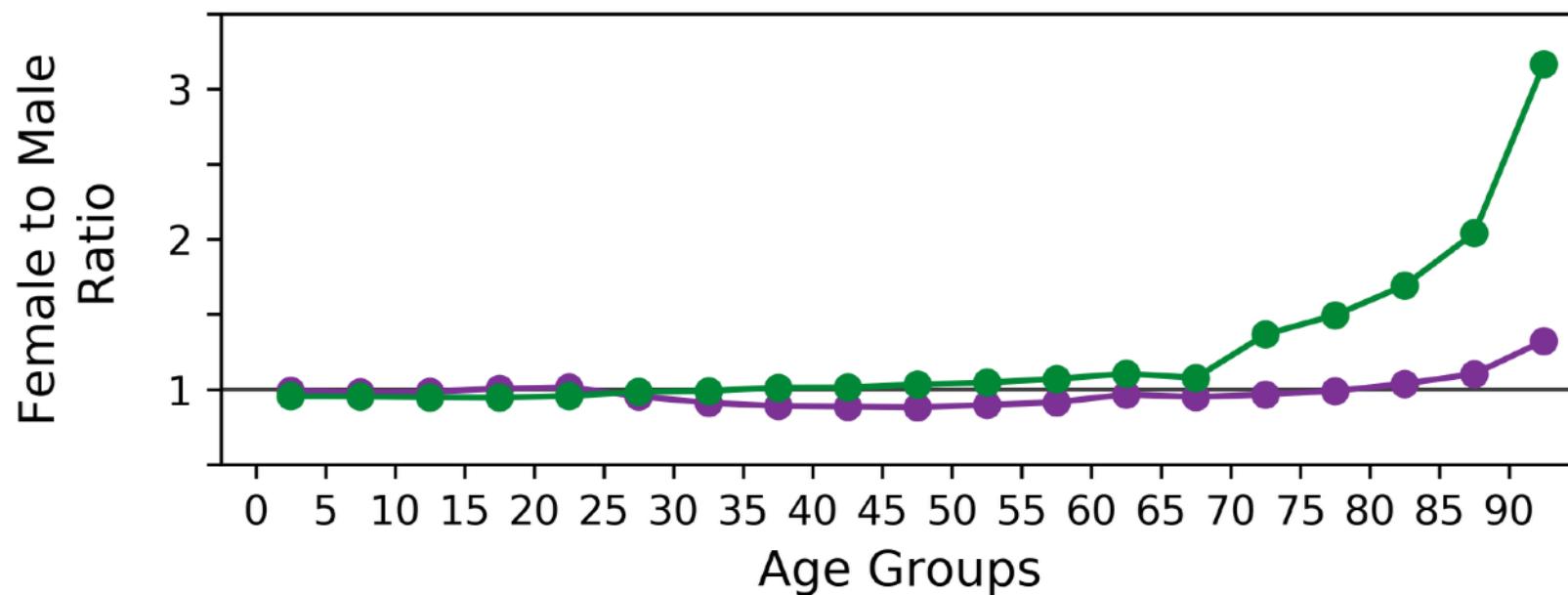
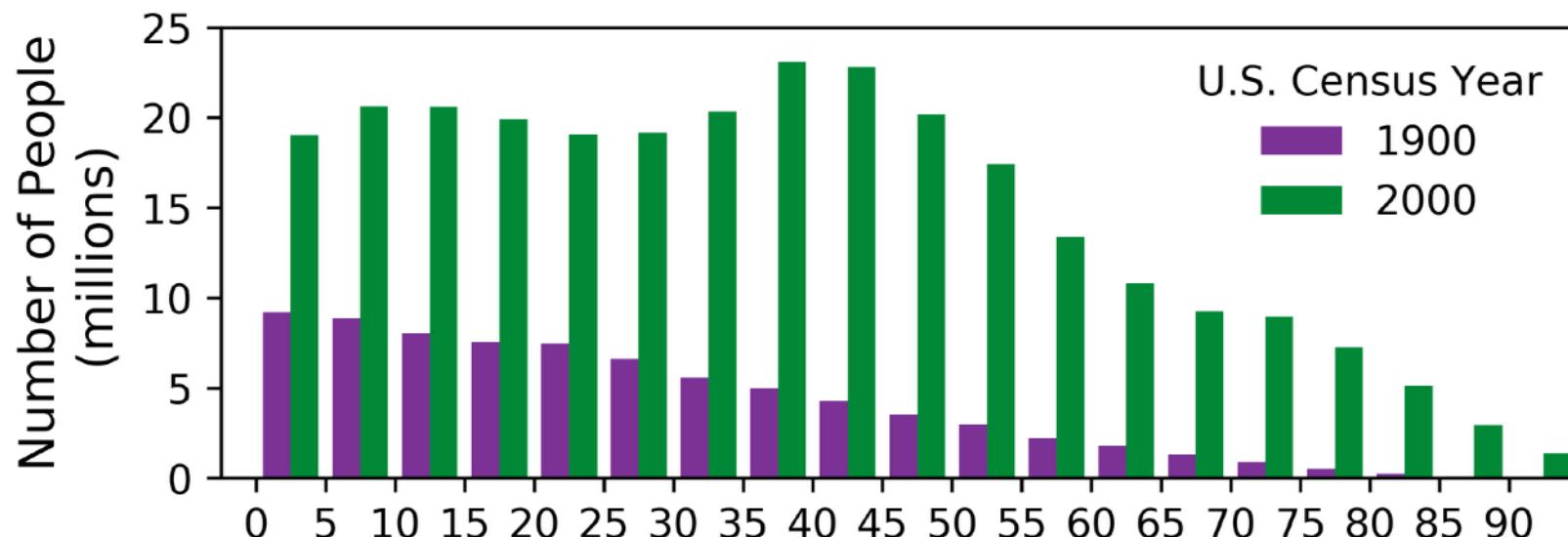
How, and why, did the distribution of ages in the U.S. change from 1900 to 2000?



How did the United States population change between
the 1900 and 2000 Census as a function of age and gender?

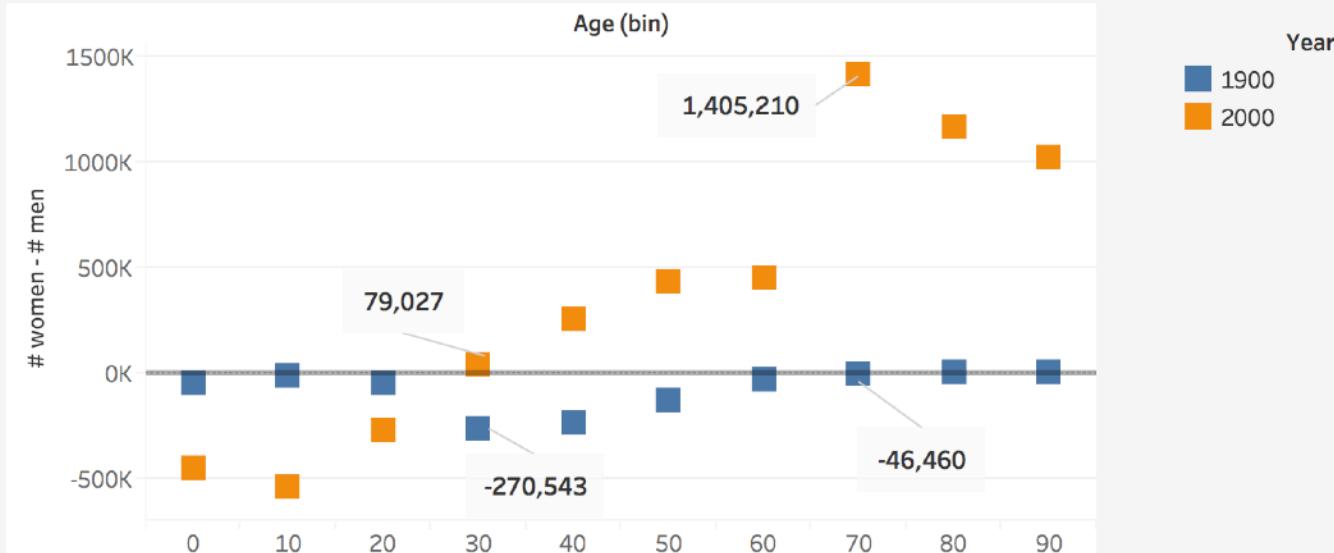


How does the female to male ratio change with total population across age groups in the U.S. for census years 1900 and 2000?

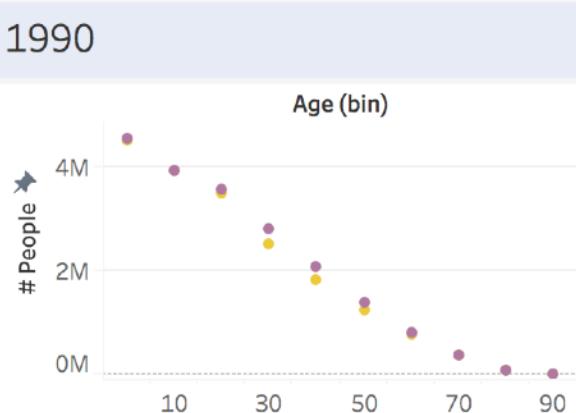


Has the 20th century affected female and male populations differently?

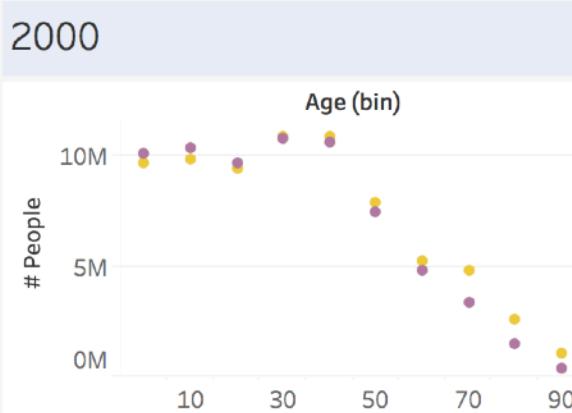
Difference between # of women and # of men



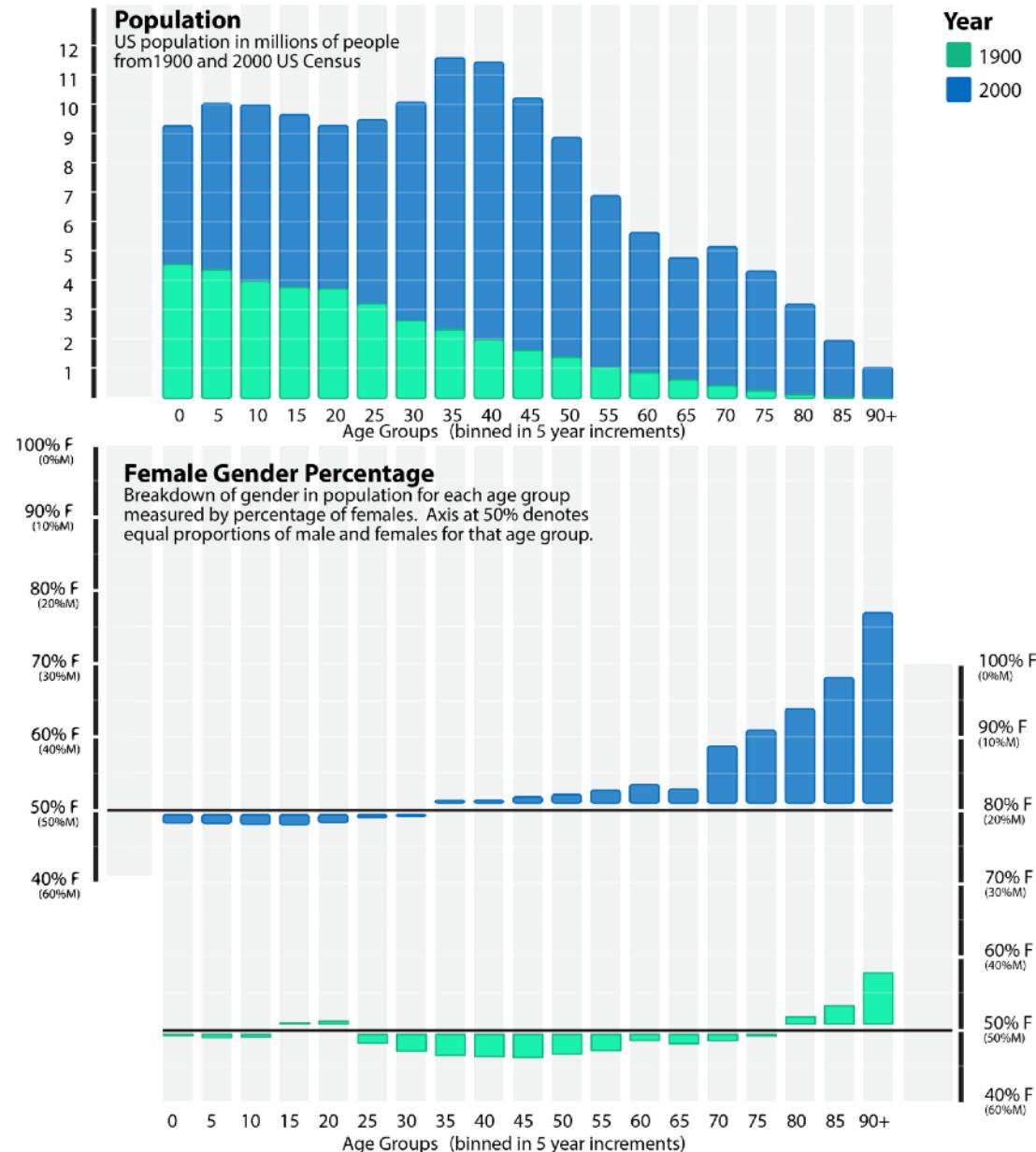
1990



2000



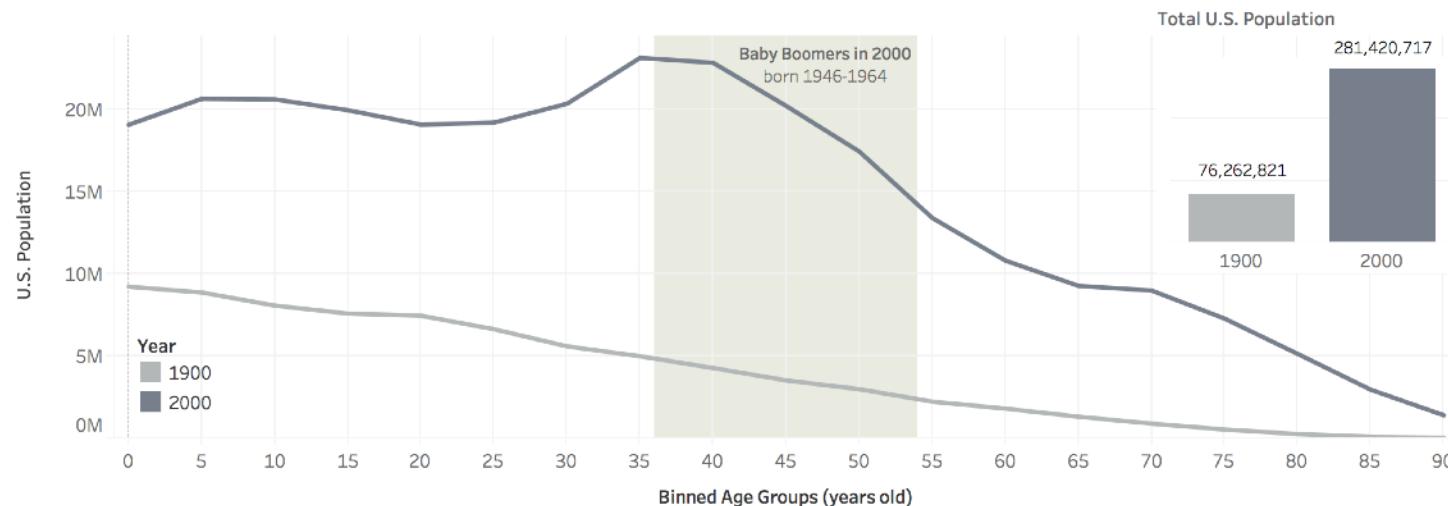
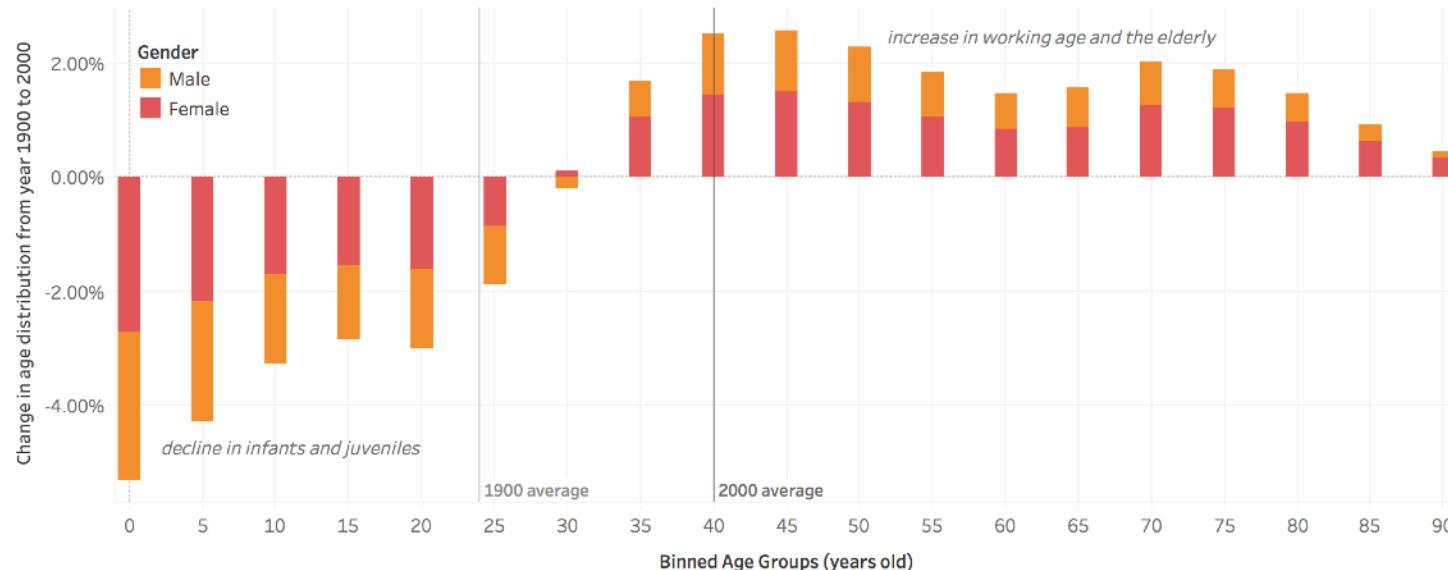
US Population and Gender Trends Across Age Group



Population data taken from the 1900 and 2000 US census.

Aging U.S. Population between 1900 and 2000

The U.S. population in 2000 increased 269% from the 1900 census, however the exceptional growth between these two periods was not evenly distributed among age groups. The change in U.S. age demographics is likely the result of an aging Baby Boomer generation (those born between 1946 and 1964).



Administrivia

A2: Exploratory Data Analysis

Use visualization software to form & answer questions

First steps:

Step 1: Pick domain & data

Step 2: Pose questions

Step 3: Profile the data

Iterate as needed

Create visualizations

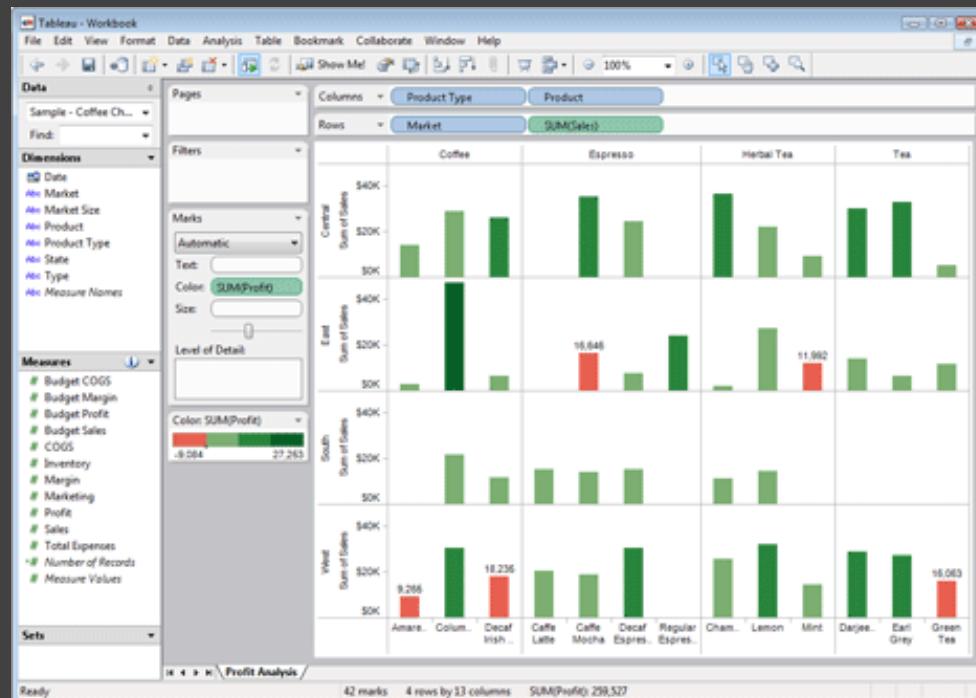
Interact with data

Refine your questions

Author a report

Screenshots of most insightful views (10+)

Include titles and captions for each view



Due by 11:59pm
Friday, Apr 13

A2: Exploratory Data Analysis

Use visualization software to form & answer questions

First steps:

Step 1: Pick domain & data

Step 2: Pose questions

Step 3: Profile the data

Iterate as needed

Create visualizations

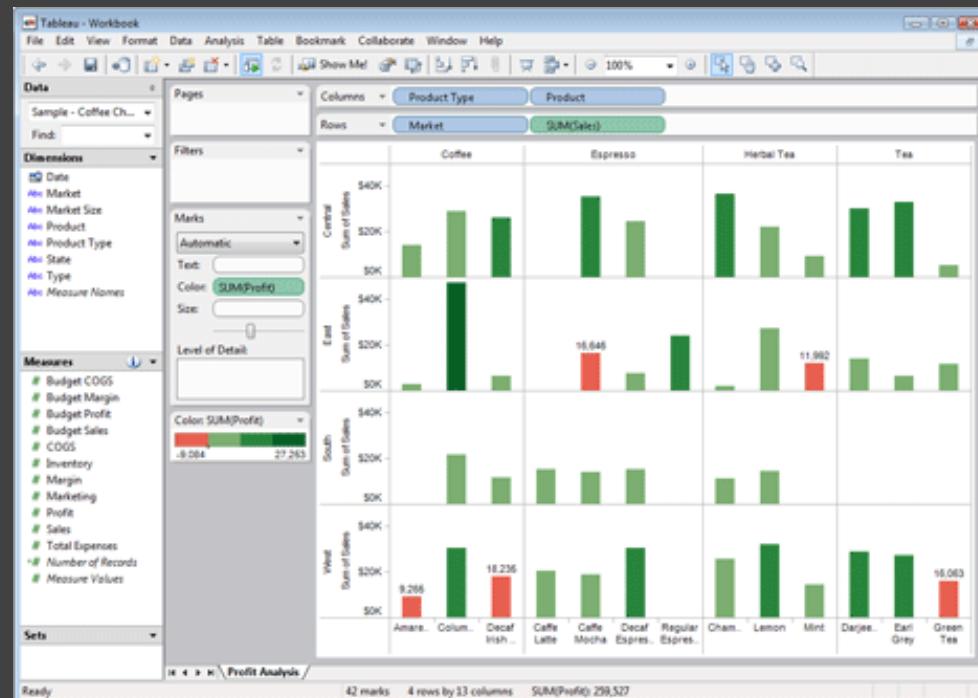
Interact with data

Refine your questions

Author a report

Screenshots of most insightful views (10+)

Include titles and captions for each view



Due by 11:59pm

Friday, Apr 13

Re-Design Exercise

Re-Design Exercise

Task: Analyze and Re-design visualization

Identify data variables (N/O/Q) and encodings

Critique the design: what works, what doesn't

Sketch a re-design to improve communication

Be ready to share your thoughts with the class

Break into groups with those sitting near you
(~4 people per group)

Effectiveness Rankings

[Mackinlay 86]

QUANTITATIVE

Position
Length
Angle
Slope
Area (Size)
Volume
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Shape

ORDINAL

Position
Density (Value)
Color Sat
Color Hue
Texture
Connection
Containment
Length
Angle
Slope
Area (Size)
Volume
Shape

NOMINAL

Position
Color Hue
Texture
Connection
Containment
Density (Value)
Color Sat
Shape
Length
Angle
Slope
Area
Volume

Teacher Salaries: Is It Really That Bad?

National and State averages for K-12 Public-School Teachers



UNITED STATES

Avg. salary: \$47,674

Avg. vacation days: 63

HOURLY

Hours per week on-site: 35.5

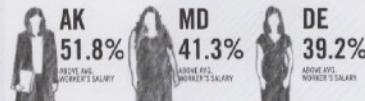
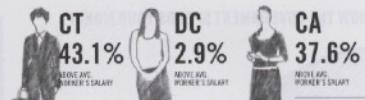
Public-School Teacher: \$34.06

Private-School Teacher: \$21.08

Average Worker: \$25.08

Police: \$22.64

Fire: \$17.91



CANADA

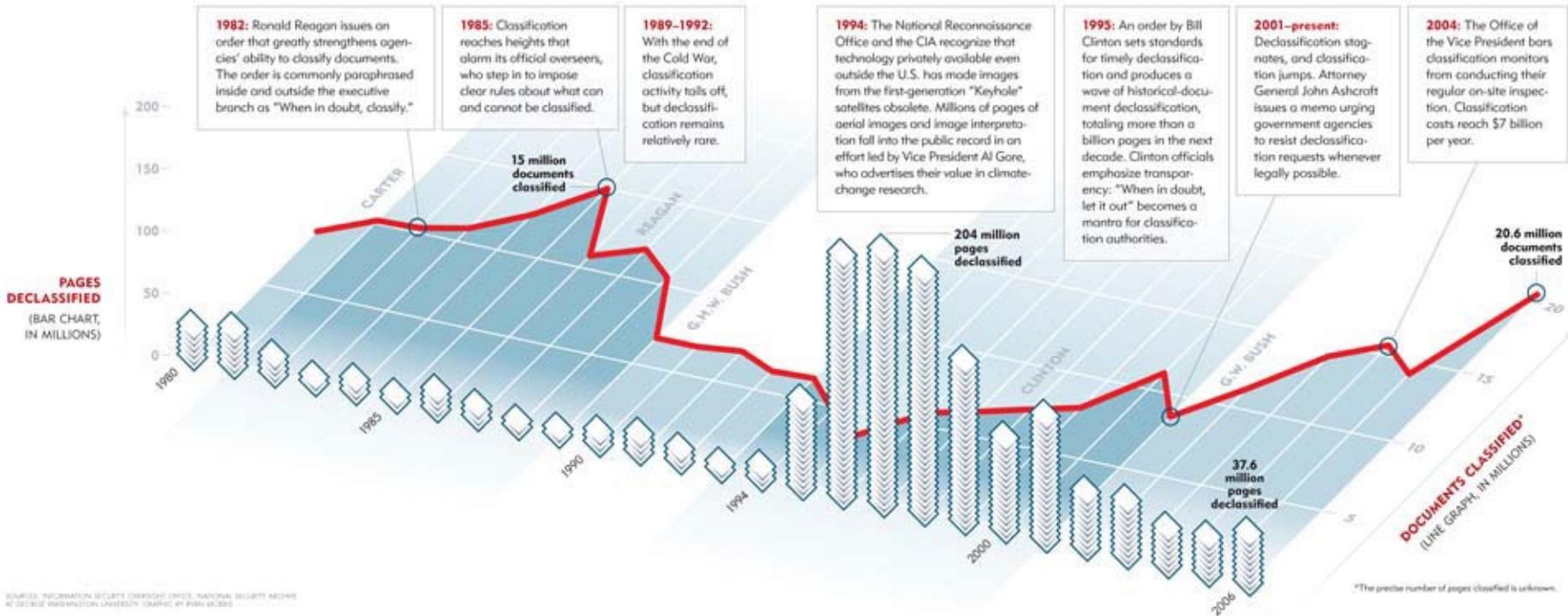
Avg. salary: \$43,000

Avg. vacation days: 50

Hourly: \$30.18

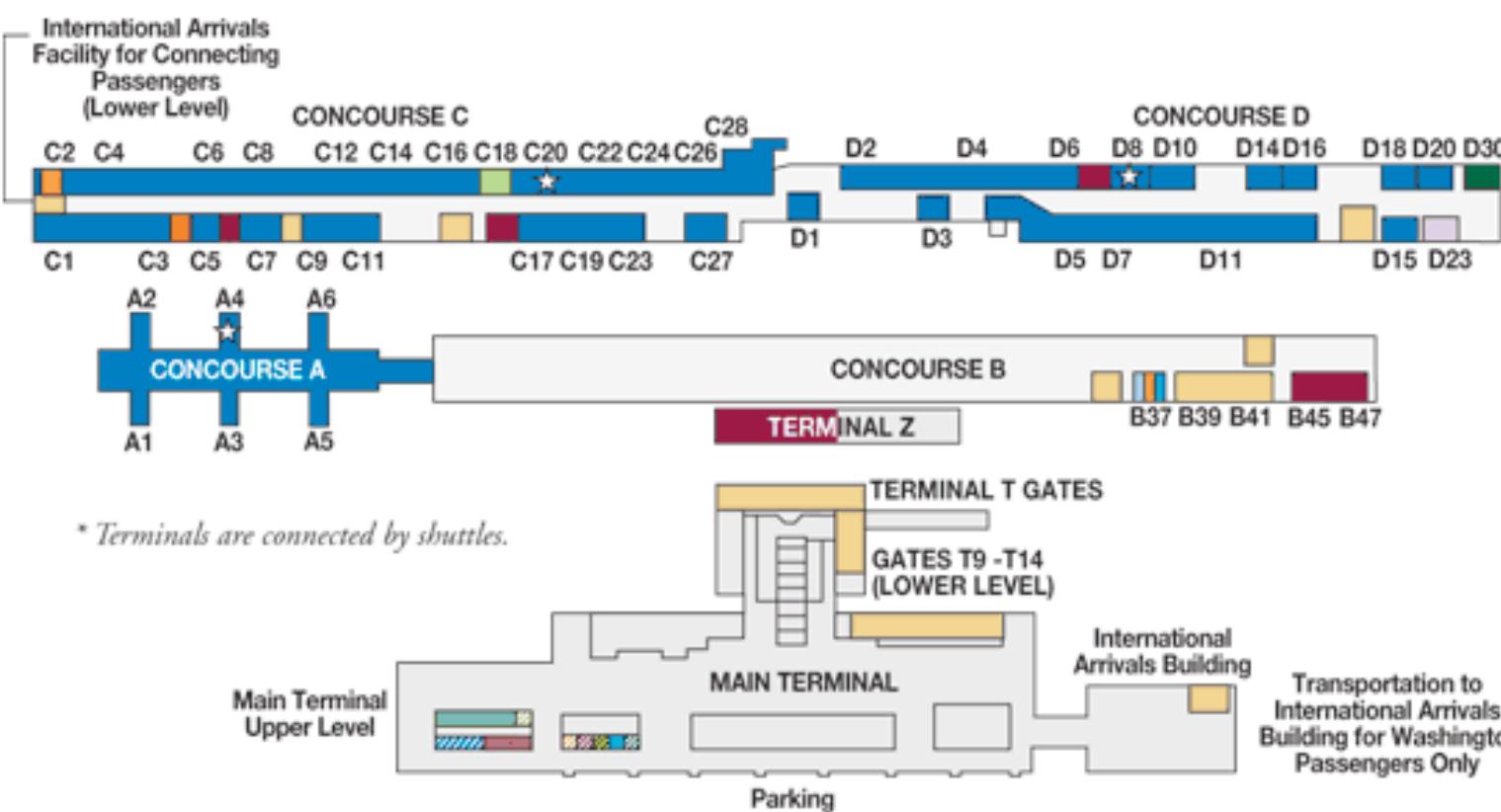
Hours per week on-site: 55.6





Source: *The Atlantic* 300 no. 2 (September 2007)

Number of Classified U.S. Documents



■ United / TED Gate Area	■ Lufthansa Check-in	■ Austrian Airlines Gate Area
■ United Premier Check-in	■ Air Canada Gate Area	■ SAS Gate Area
■ United Check-in	■ Air Canada Check-in	■ BWIA Gate
■ United International Check-in	■ Mobile Lounge Dock	■ South African Airways
■ United Red Carpet Club	■ ANA Check-in	■ US Airways Gates
■ United First International Lounge	■ ANA Fuji Lounge/Gate Area	★ United EasyCheck-in
■ Lufthansa Gate Area	■ Austrian Airlines Check-in	■ US Airways Check-in

EasyCheck-in is available at this airport.

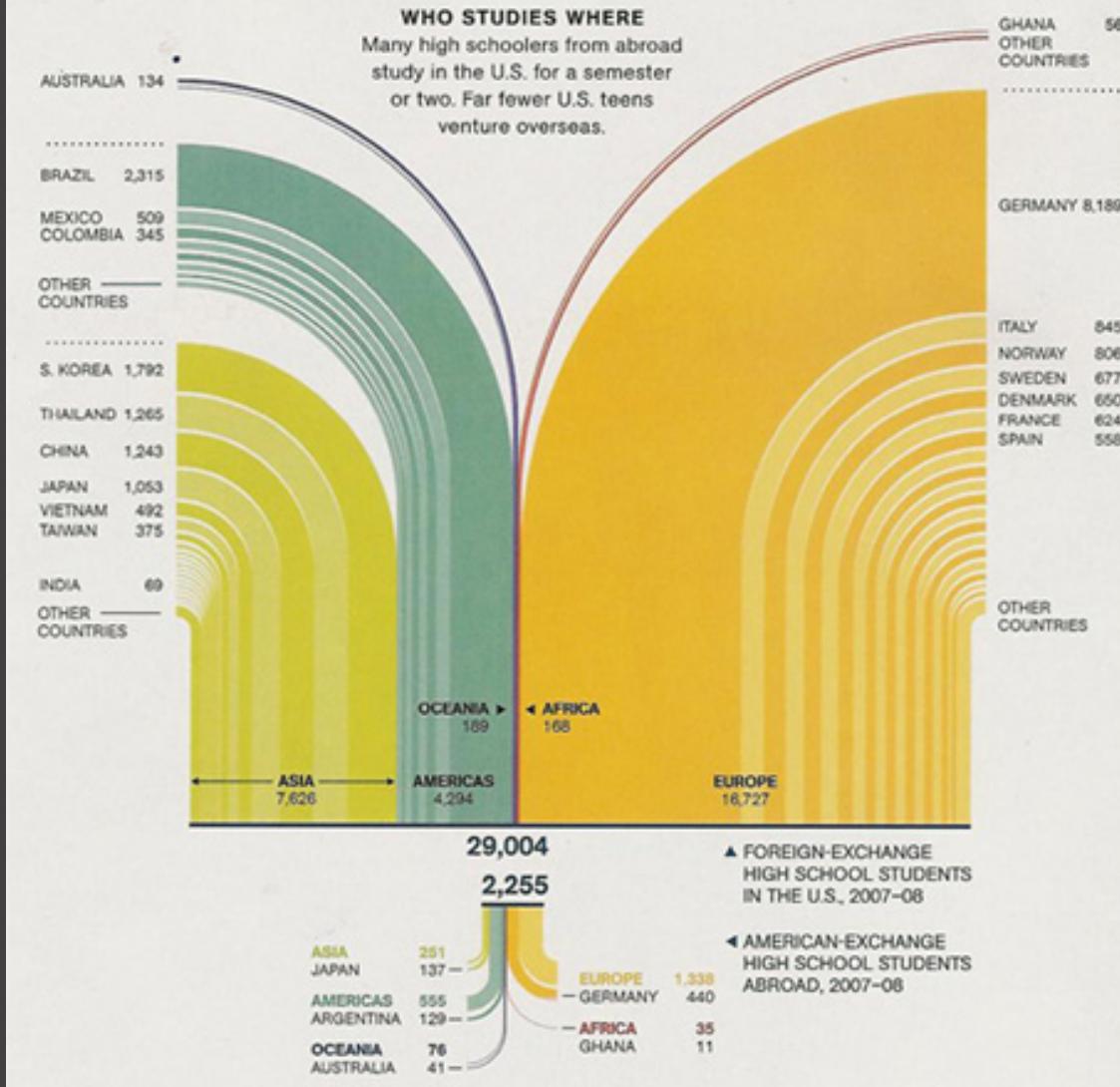


11/2006

Washington Dulles Airport Map

Source: United Airlines *Hemispheres*

GEOGRAPHY

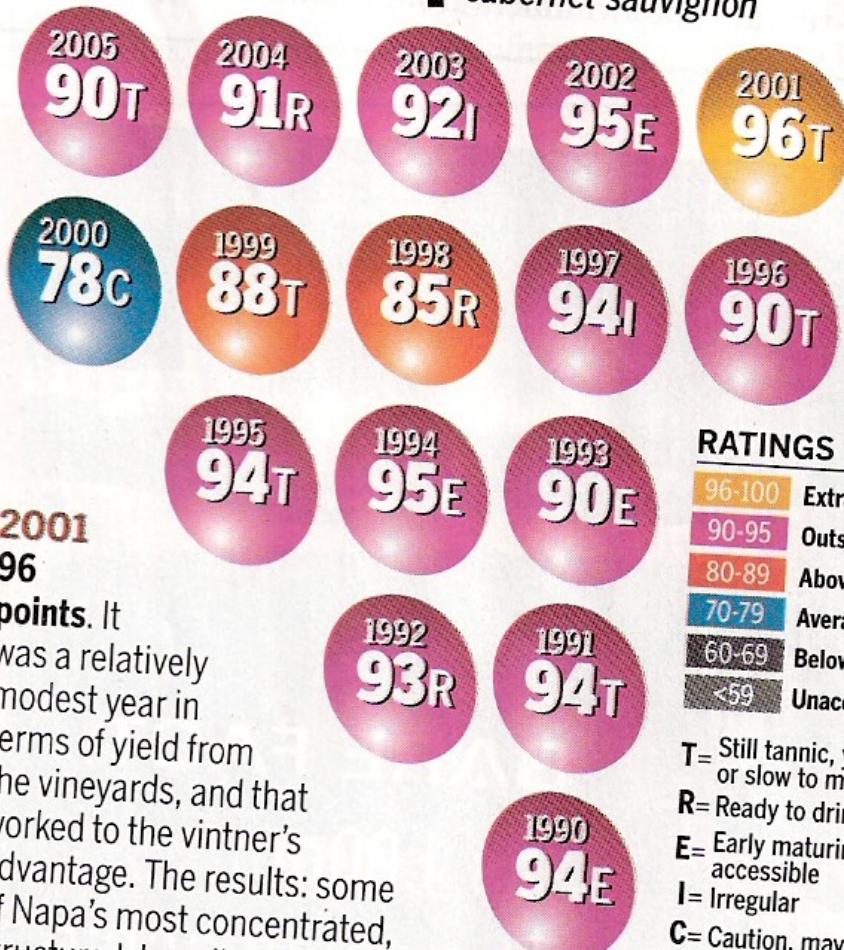


Source: *National Geographic*, September, 2008, p. 22.
 Silver, Mark. "High School Give-and-Take."

they are shaping up.

IT WAS A VERY GOOD YEAR?

Robert Parker's ratings for vintages of Napa Valley cabernet sauvignon



2001
96
points. It
was a relatively
modest year in
terms of yield from
the vineyards, and that
worked to the vintner's
advantage. The results: some
of Napa's most concentrated,
structured, long-lived wines. Built for
aging, they are rich, densely colored,

RATINGS

96-100	Extraordinary
90-95	Outstanding
80-89	Above average
70-79	Average
60-69	Below average
<59	Unacceptable

T= Still tannic, youthful,
or slow to mature

R= Ready to drink

E= Early maturing and
accessible

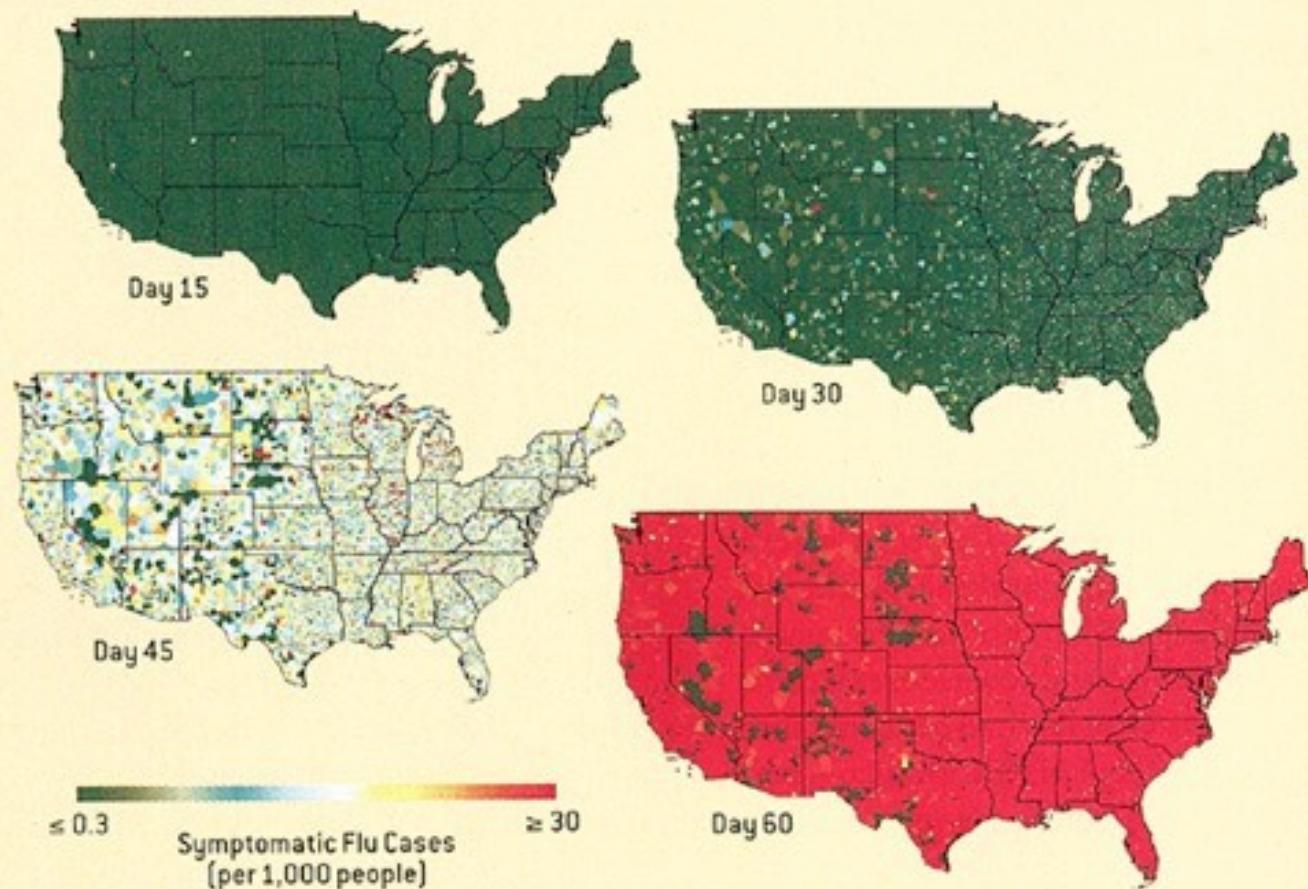
I= Irregular

C= Caution, may be too old

Source: Business Week, June 18, 2007

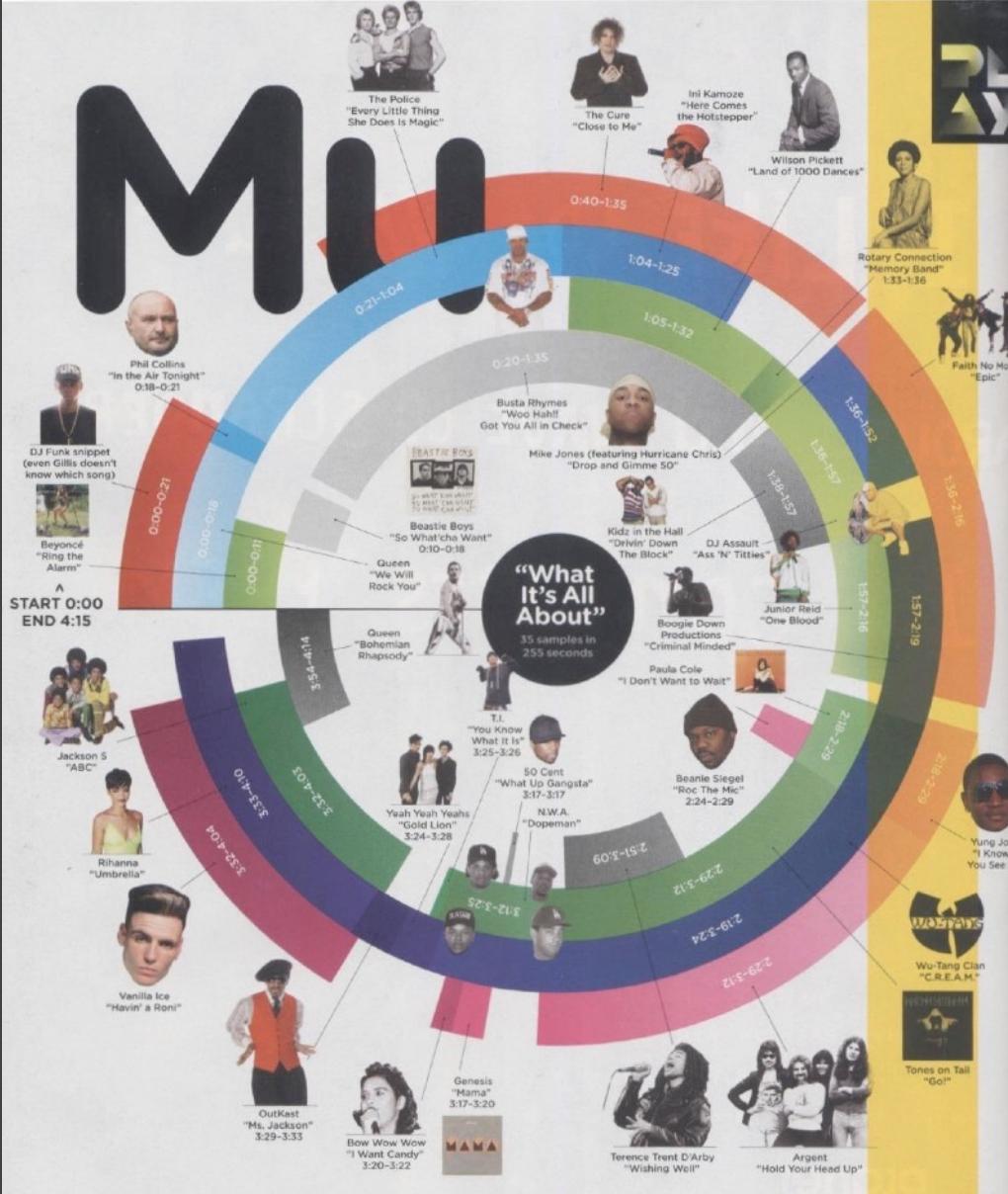
Pandemic Flu Hits the U.S.

A simulation created by researchers from Los Alamos National Laboratory and Emory University shows the first wave of a pandemic spreading rapidly with no vaccine or antiviral drugs employed to slow it down. Colors represent the number of symptomatic flu cases per 1,000 people [see scale]. Starting with 40 infected people on the first day, nationwide cases peak around day 60, and the wave subsides after four months with 33 percent of the population having become sick. The scientists are also modeling potential interventions with drugs and vaccines to learn if travel restrictions, quarantines and other disruptive disease-control strategies could be avoided.



Preparing for a Pandemic

Source: *Scientific American*, 293(5). November, 2005, p. 50



Source: *Wired Magazine*, September 2008 Edition
Music: Super Cuts (page 92)