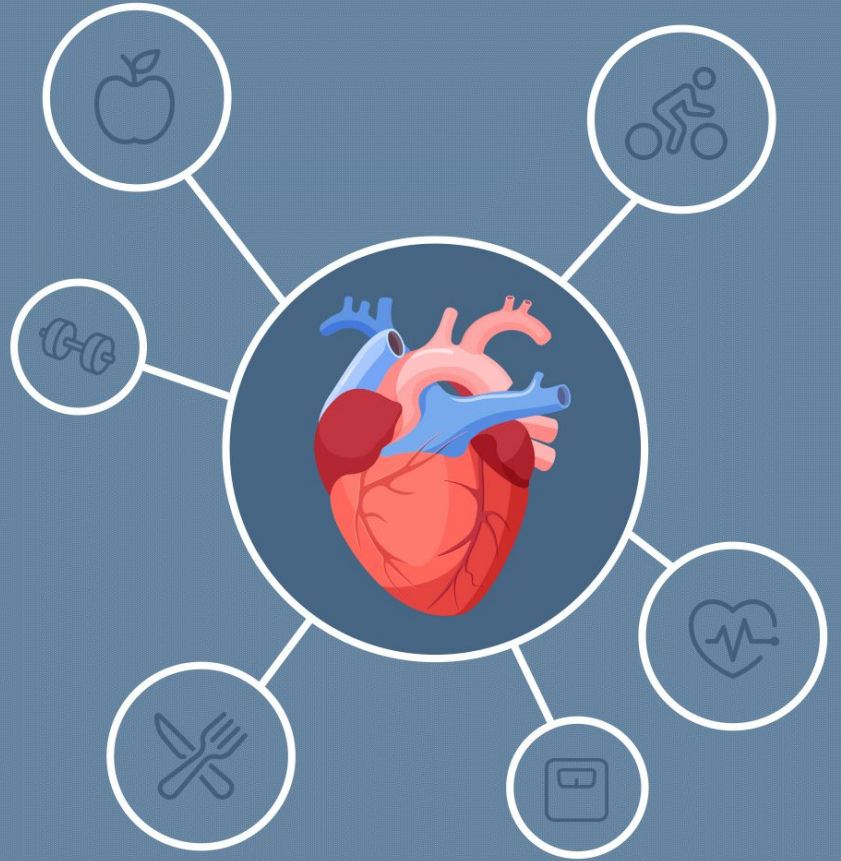


What increases risk of cardiovascular disease?

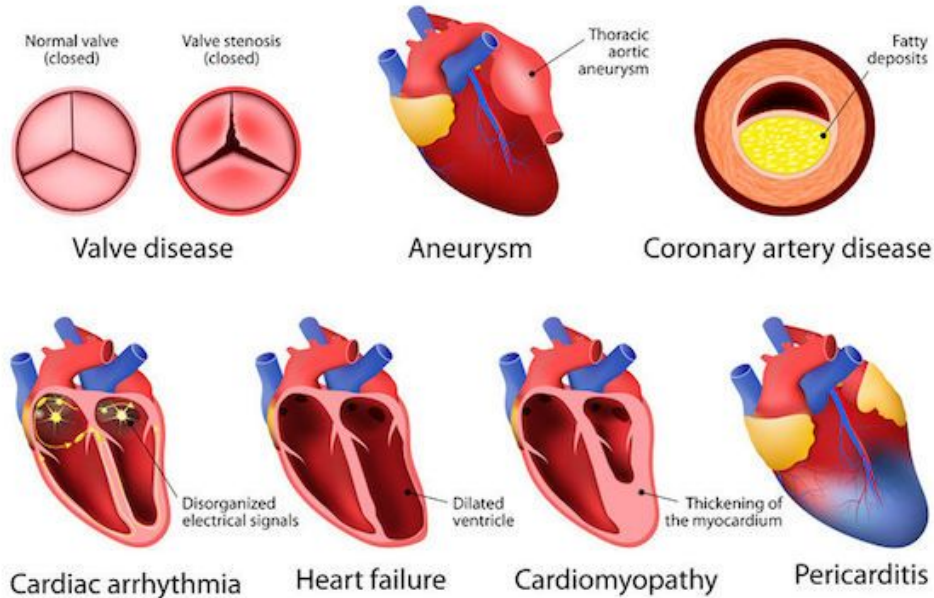
By: Jana Osea [1]

[1] University of Calgary - Department of Mathematics and Statistics



What is cardiovascular disease?

Types of heart disease

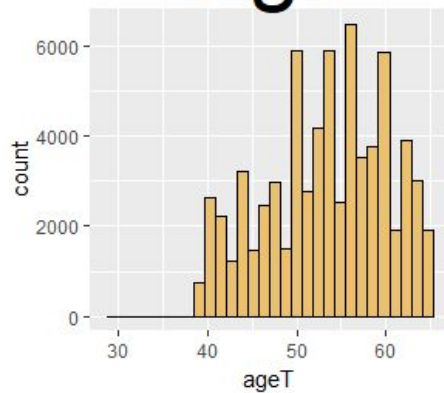


- Leading cause of death globally
- 17.9 million deaths (32.1%) in 2015
- Affects older adults more
71% from 60-80 years old
85% over 80 years old
- Men diagnosed earlier
- 90% of CVD are preventable!

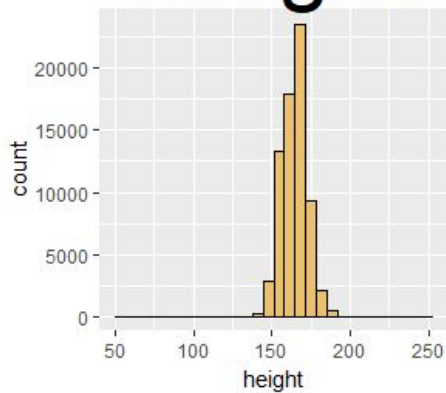
My Goal

- Quantify significance of the preventable lifestyle risks through logistic regression
- Potentially find some interesting findings that might support or contradict literature

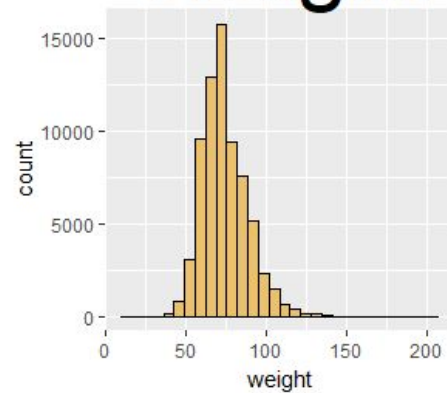
Age



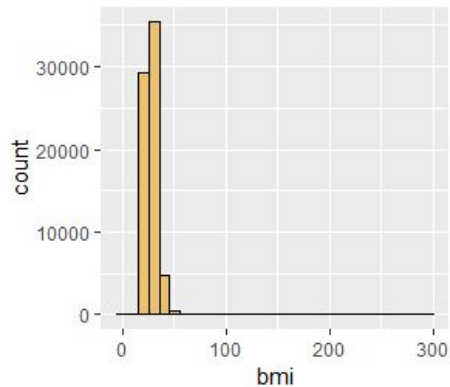
Height



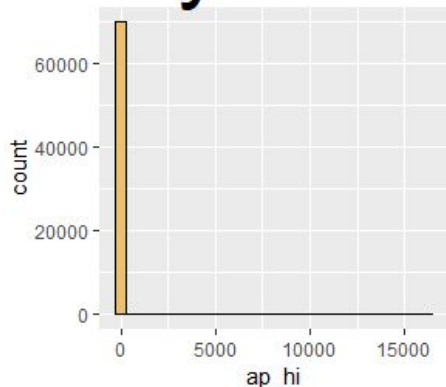
Weight



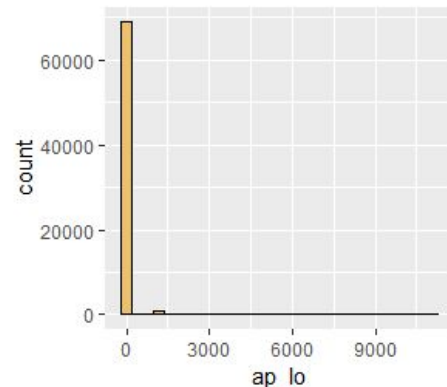
BMI

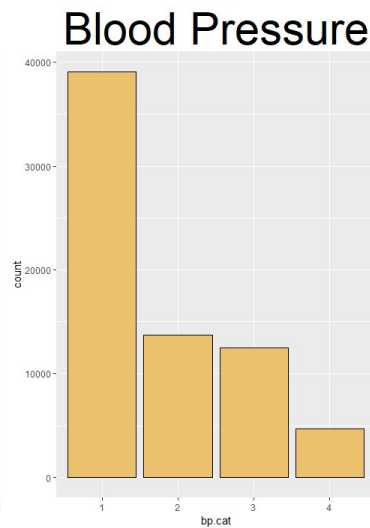
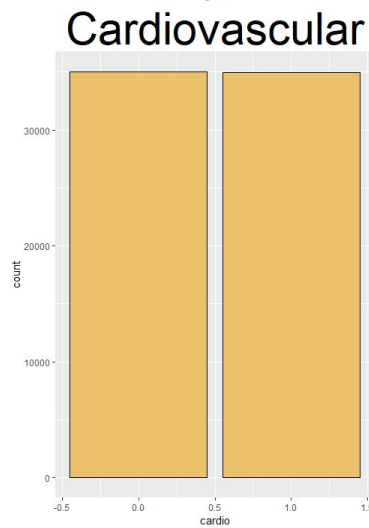
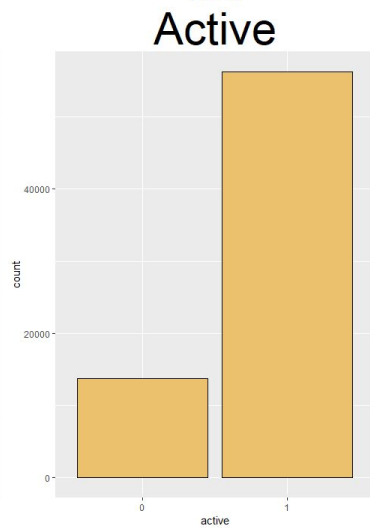
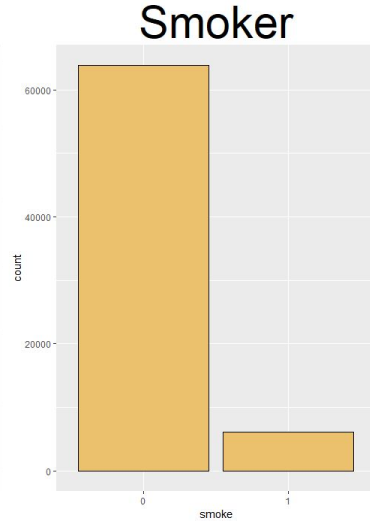
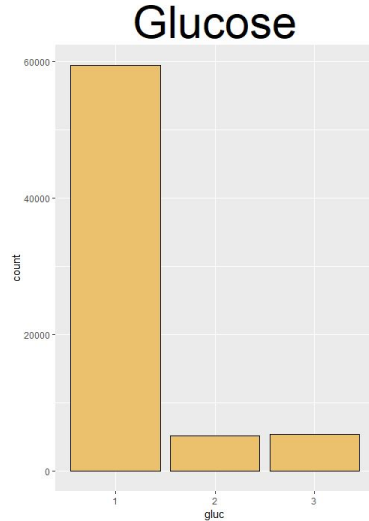
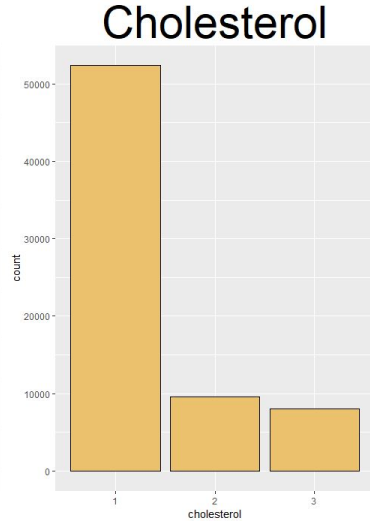
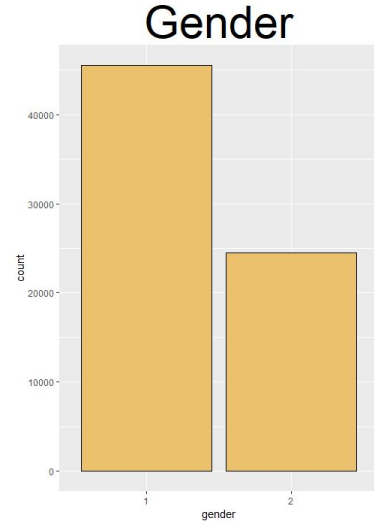


Systolic



Diastolic





```
# model 1  
m1 <- glm(cardio~.-id-weight-height-age, data=d.train)
```



```
# model 2  
m2 <- glm(formula = cardio ~ gender + cholesterol + gluc + smoke +  
          alco + active + bmi + ageT + bp.cat, data = d.train)
```

```
> vif(m2)
```

	GVIF	Df	GVIF ^{1/(2*Df)}
gender	1.149831	1	1.072302
cholesterol	1.503921	2	1.107404
gluc	1.426141	2	1.092800
smoke	1.247737	1	1.117022
alco	1.140518	1	1.067950
active	1.001781	1	1.000890
bmi	1.089641	1	1.043859
ageT	1.060740	1	1.029922
bp.cat	1.146975	3	1.023118

Methods

```
call:
glm(formula = cardio ~ gender + cholesterol + gluc + smoke +
     alco + active + bmi + ageT + bp.cat, data = d.train)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5379	-0.3272	-0.1038	0.3477	0.9382

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.3585091	0.0186281	-19.246	< 2e-16	***
gender2	0.0149534	0.0044676	3.347	0.000817	***
cholesterol2	0.0665283	0.0061173	10.876	< 2e-16	***
cholesterol3	0.1863084	0.0073568	25.324	< 2e-16	***
gluc2	0.0104220	0.0078889	1.321	0.186472	
gluc3	-0.0419284	0.0085707	-4.892	1.00e-06	***
smoke1	-0.0269633	0.0078934	-3.416	0.000636	***
alco1	-0.0509129	0.0094506	-5.387	7.19e-08	***
active1	-0.0424088	0.0049959	-8.489	< 2e-16	***
bmi	0.0040009	0.0003387	11.813	< 2e-16	***
ageT	0.0111733	0.0003027	36.908	< 2e-16	***
bp.cat2	0.2776213	0.0053875	51.530	< 2e-16	***
bp.cat3	0.4036244	0.0055811	72.320	< 2e-16	***
bp.cat4	0.4542964	0.0083502	54.405	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.1929989)

Null deviance: 12247.7 on 48990 degrees of freedom
Residual deviance: 9452.5 on 48977 degrees of freedom
AIC: 58453

Number of Fisher scoring iterations: 2

Results


```
call:
glm(formula = cardio ~ gender + cholesterol + gluc + smoke +
     alco + active + bmi + ageT + bp.cat, data = d.train)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5379	-0.3272	-0.1038	0.3477	0.9382

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.3585091	0.0186281	-19.246	< 2e-16	***
gender2	0.0149534	0.0044676	3.347	0.000817	***
cholesterol2	0.0665283	0.0061173	10.876	< 2e-16	***
cholesterol3	0.1863084	0.0073568	25.324	< 2e-16	***
gluc2	0.0104220	0.0078889	1.321	0.186472	
gluc3	-0.0419284	0.0085707	-4.892	1.00e-06	***
smoke1	-0.0269633	0.0078934	-3.416	0.000636	***
alco1	-0.0509129	0.0094506	-5.387	7.19e-08	***
active1	-0.0424088	0.0049959	-8.489	< 2e-16	***
bmi	0.0040009	0.0003387	11.813	< 2e-16	***
ageT	0.0111733	0.0003027	36.908	< 2e-16	***
bp.cat2	0.2776213	0.0053875	51.530	< 2e-16	***
bp.cat3	0.4036244	0.0055811	72.320	< 2e-16	***
bp.cat4	0.4542964	0.0083502	54.405	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.1929989)

Null deviance: 12247.7 on 48990 degrees of freedom
Residual deviance: 9452.5 on 48977 degrees of freedom
AIC: 58453

Number of Fisher scoring iterations: 2

Results


```
call:
glm(formula = cardio ~ gender + cholesterol + gluc + smoke +
     alco + active + bmi + ageT + bp.cat, data = d.train)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5379	-0.3272	-0.1038	0.3477	0.9382

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.3585091	0.0186281	-19.246	< 2e-16	***
gender2	0.0149534	0.0044676	3.347	0.000817	***
cholesterol2	0.0665283	0.0061173	10.876	< 2e-16	***
cholesterol3	0.1863084	0.0073568	25.324	< 2e-16	***
gluc2	0.0104220	0.0078889	1.321	0.186472	
gluc3	-0.0419284	0.0085707	-4.892	1.00e-06	***
smoke1	-0.0269633	0.0078934	-3.416	0.000636	***
alco1	-0.0509129	0.0094506	-5.387	7.19e-08	***
active1	-0.0424088	0.0049959	-8.489	< 2e-16	***
bmi	0.0040009	0.0003387	11.813	< 2e-16	***
ageT	0.0111733	0.0003027	36.908	< 2e-16	***
bp.cat2	0.2776213	0.0053875	51.530	< 2e-16	***
bp.cat3	0.4036244	0.0055811	72.320	< 2e-16	***
bp.cat4	0.4542964	0.0083502	54.405	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.1929989)

Null deviance: 12247.7 on 48990 degrees of freedom
Residual deviance: 9452.5 on 48977 degrees of freedom
AIC: 58453

Number of Fisher scoring iterations: 2

Results

Predictor	Log Odds	Odds
Baseline	0.318509	1.375076
Male	0.334074	1.396647
Cholesterol: Above Normal	0.388564	1.474861
Cholesterol: Well Normal	0.508106	1.66214
Glucose: Well Above Normal	0.266942	1.305965
Glucose: Above Normal	0.388564	1.474861
Smoker: No	0.286537	1.331808
Alcohol: No	0.27335	1.31436
Active: No	0.274875	1.316365
Blood Pressure: Elevated	0.59456	1.812234
Blood Pressure: High	0.726219	2.06725
Blood Pressure: Extremely High	0.771334	2.162649

Results

Conclusion

- Literature is accurate!
- Sex, Cholesterol, Glucose, Smoking Status, Alcohol, Blood pressure are all factors that can indicate for increased odds for Cardiovascular Disease

References

- Mendis S, Puska P, Norrving B (2011). *Global Atlas on Cardiovascular Disease Prevention and Control* (PDF). World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization. pp. 3–18. ISBN 978-92-4-156437-3. Archived (PDF) from the original on 2014-08-17.
- Wang H, Naghavi M, Allen C, Barber RM, Bhutta ZA, Carter A, et al. (GBD 2015 Mortality and Causes of Death Collaborators) (October 2016). "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015". *Lancet*. 388 (10053): 1459–1544. doi:10.1016/S0140-6736(16)31012-1. PMC 5388903. PMID 27733281.
- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, et al. (January 2013). "Heart disease and stroke statistics--2013 update: a report from the American Heart Association". *Circulation*. 127 (1): e6–e245. doi:10.1161/cir.0b013e31828124ad. PMC 5408511. PMID 23239837.
- Mendis S, Puska P, Norrving B (2011). *Global atlas on cardiovascular disease prevention and control* (1 ed.). Geneva: World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization. p. 48. ISBN 978-92-4-156437-3.

Thank you!
Questions?