

# Properties that influence wine score

Group 28: Aishwin Tikku, Mengran Li, Steven Kwok, Shaoquan Li, Shuning Li

## Introduction {#sec : Intro}

The whole data set has 7 factors and 2000 observations.

## Exploratory Data Analysis { #sec : EDA}



Figure 1: Histogram of Price and Points

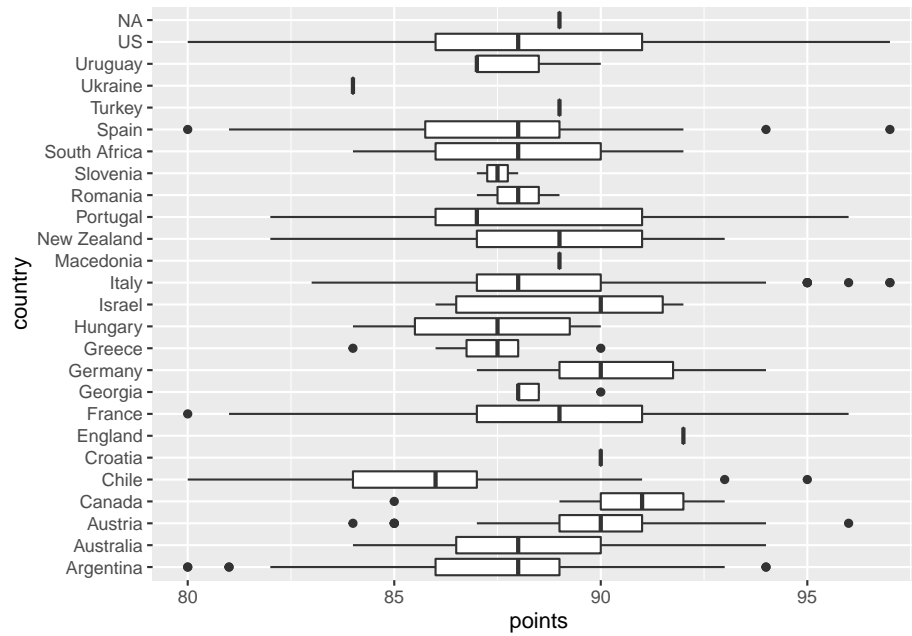


Figure 2: Boxplot of points by country

#### Pearson's Chi-squared test

data: Data\$country and Data\$score  
X-squared = 58.669, df = 24, p-value = 9.821e-05

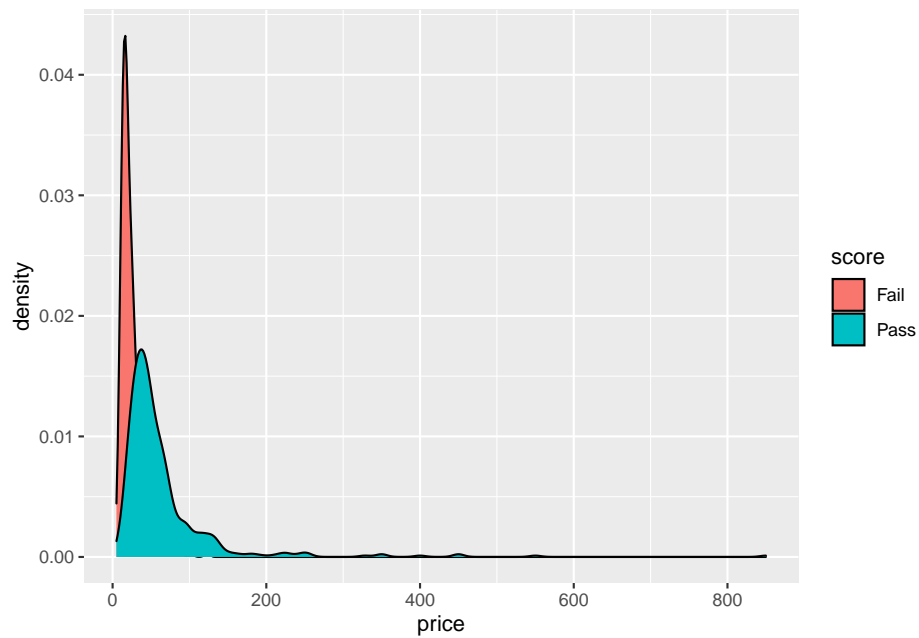


Figure 3: Density plot by score

Table 1: Cross Table of Country and Score

	Fail	Pass
Argentina	0.7826087	0.2173913
Australia	0.7826087	0.2173913
Austria	0.6000000	0.4000000
Canada	0.2857143	0.7142857
Chile	0.9315068	0.0684932
Croatia	1.0000000	0.0000000
England	0.0000000	1.0000000
France	0.6768802	0.3231198
Georgia	1.0000000	0.0000000
Germany	0.6153846	0.3846154
Greece	1.0000000	0.0000000
Hungary	1.0000000	0.0000000
Israel	0.5714286	0.4285714
Italy	0.8020134	0.1979866
Macedonia	1.0000000	0.0000000
New Zealand	0.6521739	0.3478261
Portugal	0.7236842	0.2763158
Romania	1.0000000	0.0000000
Slovenia	1.0000000	0.0000000
South Africa	0.8076923	0.1923077
Spain	0.7857143	0.2142857
Turkey	1.0000000	0.0000000
Ukraine	1.0000000	0.0000000
Uruguay	1.0000000	0.0000000
US	0.7333333	0.2666667

Table 2: Summary Statistics of GLM Model

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-3.2480326	0.3928483	-8.2679055	0.0000000
price	0.0618039	0.0036883	16.7567457	0.0000000
countryAustralia	-0.3413567	0.7321541	-0.4662362	0.6410465
countryAustria	0.8275812	0.5284127	1.5661643	0.1173102
countryCanada	1.1123801	1.0558659	1.0535240	0.2921009
countryChile	-0.9034702	0.6626117	-1.3634986	0.1727254
countryCroatia	-14.3687014	2399.5447495	-0.0059881	0.9952222
countryEngland	17.4037507	2399.5447491	0.0072529	0.9942130
countryFrance	0.1711982	0.4120220	0.4155076	0.6777704
countryGeorgia	-14.4953783	1180.6530543	-0.0122774	0.9902043
countryGermany	0.3629247	0.6422801	0.5650568	0.5720351
countryGreece	-14.8607632	819.5126061	-0.0181337	0.9855322
countryHungary	-16.4098955	1034.4621463	-0.0158632	0.9873435
countryIsrael	0.5012816	0.9311403	0.5383524	0.5903338
countryItaly	-1.0063838	0.4326208	-2.3262494	0.0200052
countryMacedonia	-14.5541129	2399.5447493	-0.0060654	0.9951606
countryNew Zealand	1.1326296	0.6096573	1.8578135	0.0631955
countryPortugal	0.5513142	0.5130956	1.0744864	0.2826048
countryRomania	-14.2089305	1678.1159242	-0.0084672	0.9932442
countrySlovenia	-14.8496606	1680.7737693	-0.0088350	0.9929508
countrySouth Africa	0.3572279	0.6492801	0.5501907	0.5821886
countrySpain	0.2386346	0.4926448	0.4843948	0.6281057
countryTurkey	-14.2450936	2399.5447496	-0.0059366	0.9952633
countryUkraine	-13.8742705	2399.5447502	-0.0057820	0.9953866
countryUruguay	-14.7201546	1385.0392851	-0.0106280	0.9915203
countryUS	-0.0668919	0.3886999	-0.1720915	0.8633656

## Formal Data Analysis

NewZealand and Italy are significantly different from others

GLM MOdel for New Zealand

GLM Model for Italy

GLM Model & Predicted Values Plot based on only Price

\$price

Table 3: GLM Model for New Zealand

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-25.9450573	1.773037e+04	-0.0014633	0.9988324
price	0.3585993	2.112316e-01	1.6976590	0.0895721
provinceCentral Otago	59.7177350	1.080974e+05	0.0005524	0.9995592
provinceHawke's Bay	-2.8687940	2.507453e+04	-0.0001144	0.9999087
provinceKumeu	30.3741594	2.507453e+04	0.0012114	0.9990335
provinceMarlborough	17.7017003	1.773037e+04	0.0009984	0.9992034
provinceMartinborough	36.9340816	1.929623e+04	0.0019141	0.9984728
provinceWairau Valley	-1.0757978	2.507453e+04	-0.0000429	0.9999658
varietyMerlot	31.4499572	2.507453e+04	0.0012543	0.9989992
varietyPinot Noir	-22.1715905	1.051490e+05	-0.0002109	0.9998318
varietyRiesling	-36.0322304	1.362181e+04	-0.0026452	0.9978894
varietyRosé	-18.4188988	1.773037e+04	-0.0010388	0.9991711

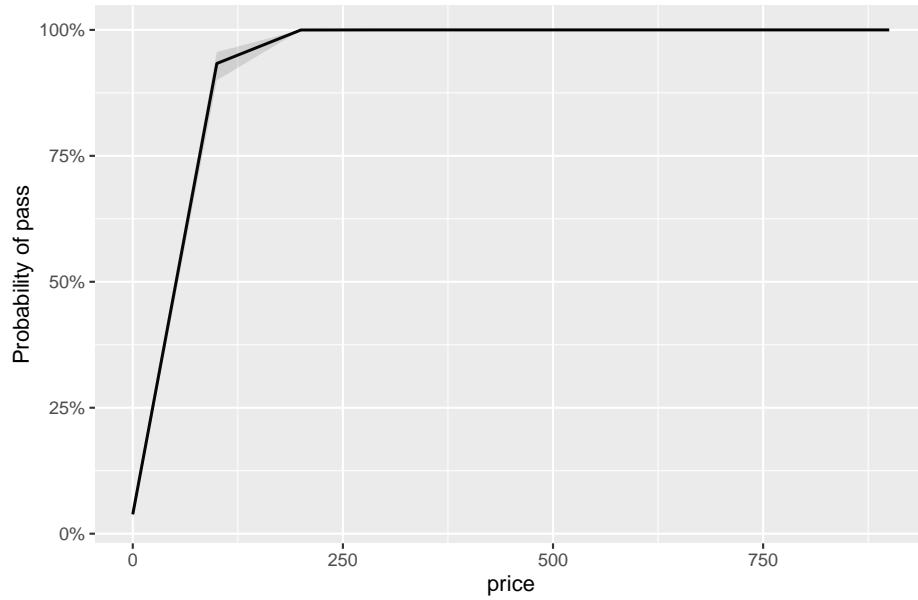


Figure 4: Pass Probability Plot by Price

Table 4: GLM Model for Italy

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-5.4913621	9.776996e+03	-0.0005617	0.9995519
price	0.0702738	1.350710e-02	5.2027307	0.0000002
provinceItaly Other	-16.1990868	2.024736e+04	-0.0008001	0.9993616
provinceLombardy	15.5740188	2.231830e+04	0.0006978	0.9994432
provinceNortheastern Italy	-20.7067445	9.797513e+03	-0.0021135	0.9983137
provincePiedmont	1.7338206	9.776996e+03	0.0001773	0.9998585
provinceSicily & Sardinia	19.2911547	5.153375e+03	0.0037434	0.9970132
provinceSouthern Italy	0.8081251	9.776996e+03	0.0000827	0.9999341
provinceTuscany	16.5883448	5.153375e+03	0.0032189	0.9974317
provinceVeneto	17.5416845	5.153375e+03	0.0034039	0.9972841
varietyAlbana	-16.1288130	2.024736e+04	-0.0007966	0.9993644
varietyArneis	-18.5653713	1.773037e+04	-0.0010471	0.9991645
varietyBarbera	-18.5592672	1.015066e+04	-0.0018284	0.9985412
varietyCabernet	-16.7963905	1.773037e+04	-0.0009473	0.9992441
varietyCabernet Franc	-18.2811107	8.308573e+03	-0.0022003	0.9982444
varietyCabernet Sauvignon	-35.3299124	1.402586e+04	-0.0025189	0.9979902
varietyCarignano	5.7121693	1.958056e+04	0.0002917	0.9997672
varietyChardonnay	-37.2860731	1.245119e+04	-0.0029946	0.9976107
varietyCiliegiolo	-33.9118119	1.958056e+04	-0.0017319	0.9986181
varietyCorvina, Rondinella, Molinara	-17.0796488	8.308573e+03	-0.0020557	0.9983598
varietyDolcetto	-18.0611011	1.021325e+04	-0.0017684	0.9985890
varietyFalanghina	-16.9369381	1.253726e+04	-0.0013509	0.9989221
varietyFiano	-17.1477594	1.773037e+04	-0.0009671	0.9992283
varietyGlera	-33.9839315	9.554127e+03	-0.0035570	0.9971619
varietyGreco	-17.2883069	1.773037e+04	-0.0009751	0.9992220
varietyInzolia	-35.5761138	1.500677e+04	-0.0023707	0.9981085
varietyLambrusco	-16.6207294	2.024736e+04	-0.0008209	0.9993450
varietyMalvar	3.1724560	2.127986e+04	0.0001491	0.9998810
varietyMalvasia	-37.8795497	1.958056e+04	-0.0019345	0.9984565
varietyMerlot	-17.3864853	8.308573e+03	-0.0020926	0.9983304
varietyMontepulciano	-17.1486380	1.231903e+04	-0.0013920	0.9988893
varietyMoscato	-17.9329073	1.773037e+04	-0.0010114	0.9991930
varietyNegroamaro	2.9263927	1.955325e+00	1.4966276	0.1344902
varietyNerello Mascalese	2.7762695	1.500677e+04	0.0001850	0.9998524
varietyNero d'Avola	-35.4228083	1.148180e+04	-0.0030851	0.9975384
varietyPasserina	-15.9882654	2.024736e+04	-0.0007896	0.9993700
varietyPetit Manseng	2.6805396	2.127986e+04	0.0001260	0.9998995
varietyPinot Grigio	4.2207980	1.289648e+04	0.0003273	0.9997389
varietyPinot Nero	5.0630385	1.039383e+04	0.0004871	0.9996113
varietyPrimitivo	-16.7261168	1.773037e+04	-0.0009434	0.9992473
varietyProsecco	-34.4431765	1.496564e+04	-0.0023015	0.9981637
varietyPrugnolo Gentile	-34.4740021	1.958056e+04	-0.0017606	0.9985952
varietyRed Blend	-16.7682577	8.308573e+03	-0.0020182	0.9983897
varietyRibolla Gialla	3.3710544	1.704139e+04	0.0001978	0.9998422
varietyRiesling	4.7184790	2.127986e+04	0.0002217	0.9998231
varietySagrantino	-20.4857869	2.024736e+04	-0.0010118	0.9991927
varietySangiovese	-15.7580040	8.308573e+03	-0.0018966	0.9984867
varietySangiovese Grosso	-15.3646591	8.308573e+03	-0.0018493	0.9985245
varietySauvignon	3.9830417	1.558367e+04	0.0002556	0.9997961
varietySparkling Blend	-33.8110450	1.504046e+04	-0.0022480	0.9982064
varietySyrah	-16.3357565	8.308573e+03	-0.0019661	0.9984313
varietyTocai	4.9298003	2.127986e+04	0.0002316	0.9998152
varietyTrebiano	-15.7071703	2.024736e+04	-0.0007758	0.9993810
varietyTurbiana	-32.5873765	2.479432e+04	-0.0013143	0.9989513
varietyVerdicchio	-16.3699348	1.255737e+04	-0.0013036	0.9989599