

40.016: The Analytics Edge

Week 2 Lecture 1

LINEAR REGRESSION:
PREDICTING QUALITY AND PRICE OF WINE

Term 5, 2022





<https://www.youtube.com/watch?v=CzPrvVQgF9c>

Bordeaux wines

- Excellent weather conditions conducive for growing vines.
- Difference in price and quality between years.
- Aged wine is supposed to taste good.
Younger wines not so pleasant to taste.
- Hence the importance of wine experts.
 - Robert Parker
 - <https://www.robertparker.com/>

2020 2019 2018 2017 2016 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2004 2003 2002 2001 2000 1999 1998 1997 1996 1995 1994 1993 1992 1991 1990 1989 1988 1987 1986

AUSTRIA RIESLING & GRÜNER VELTLINER	98T 93T 92R 90R 95T 78E 91R 89T 89E 88I 89I 88R 90R 91I 87I 88I 89I 89T 88I 85C 95R 82C 96R 84C 90I 87C 89C 88C 84C 94R NT NT NT NT 96R
-------------------------------------	---

FRANCE

ALSACE	NT NT 93T 90E 94T 89I 86T 87R 86E 89I 82I 89R 91R 79I 87R 86E 82R 86R 91R 90R 87R 90R 57R 87R 89R 90R 87R 85C 78C 93R 93R 88R 83C 82C
BORDEAUX: ST JULIEN/PAUILLAC/ST ESTEPHE	NT 96T 91I 97T 94T 93T 81C 92E 88E 98T 93E 91E 86E 87E 95T 88T 95T 88T 88R 96T 88R 87T 84R 96T 92T 85C 78C 79C 75R 58E 90E 87R 82R 94T
BORDEAUX: MARGAUX	NT 96T 90I 97T 94T 90T 80C 89E 87E 95T 97E 90E 86E 88E 98T 87T 88I 88T 89E 94T 89R 86T 82R 88T 88E 85C 77C 75C 74R 90E 86E 85R 76R 90T
BORDEAUX: GRAVES/PESSAC LEONGAN	NT 95T 92I 97T 95T 93T 81C 91E 86E 95T 95E 91E 87E 87E 96T 88T 88I 87T 88R 97T 88R 94T 86R 86E 89E 88E 86C 75C 74R 90R 89R 89R 84R 89E
BORDEAUX: POMEROL	NT 95T 90I 97T 96T 94T 84R 94E 88R 95T 93E 96E 86I 90T 95T 88E 84E 85E 90E 95T 88R 96T 87R 85E 92T 89T 87C 82C 58C 96R 96R 89R 85C 87T
BORDEAUX: ST EMILION	NT 94T 89I 95T 95T 92T 82R 93E 87R 94T 93E 92E 86I 88E 92T 88E 90I 87E 90E 96T 88R 96T 86R 87T 88E 86T 84C 75C 59C 98R 88R 74C 88E
BORDEAUX: BARSAC/SAUTERNES	NT 90E 91E 92E 95E 92E 92E 88E 93E 90R 97E 89R 94R 88R 96R 87R 89R 85R 93R 88E 88E 87E 89E 87E 85E 78E 70C 68C 70C 98R 90R 90R 70R 94R
BURGUNDY: COTE DE NUITS (RED)	98E 93T 93E 97T 90T 92E 92E 95T 91E 96E 95E 88I 91E 89I 90T 83C 89R 94R 91R 87R 92R 84I 89R 89T 90R 72C 85C 69C 86R 93R 85C 84C 77C 65C
BURGUNDY: COTE DE BEAUNE (RED)	96E 94T 92E 92E 96T 91E 89E 91E 90E 94E 95E 89I 90E 82I 90T 79C 87T 92B 88R 86R 93R 82C 88R 89R 85R 73C 80C 82R 77C 92R 86C 87C 79C 72C
BURGUNDY (WHITE)	93T 91E 97E 87I 94E 97E 90E 92E 91E 94E 91E 94E 91I 90E 88E 91R 84R 92R 86C 88C 89C 84C 89C 92C 93C 77C 72C 90C 70C 87C 90C 82C
BURGUNDY: BEAUJOLAIS	94E 93E 95E 95I 90I 93R 86R 90R 91R 93R 97B 86R 85R 89R 95R 81C 93R 86C 75C 91R 89R 84C 87C 82C 87C 85C 80C 77C 88C 86C 92C 86C 85C 84C
CHAMPAGNE	NT 92E 87I 94T 93E 94E 95T 96T 87R 89R 92R 99T 80R 86R 88I 90T 88I 95T 88R 92R 92R 93R 90I 97T 95T NT 88E NT NT 93R 95R 95R NT NT
JURA	94E 93E 91I 93T 94T 93E 90I 93E 91R 94E 90R 91T 91I 90E 96R 92R 94R 93R 85R NT
LANGUEDOC	NT 92T 91E 90I 91T 92E 93T 87E 88E 88I 91R 94E 91R 87R 93R 90I 88R 88C 87I NT
ROUSSILLON	NT 91I 89I 91T 91T 92E 93T 88E 94E 89R 91R 94E 91R 87R 92R 90I 88R 88C 87I NT
LOIRE VALLEY (WHITE)	96T 95E 94T 93R 92T 91R 80I 83E 86E 92T 88I 90I 84I 83E 94E 82C 82R 96R 82C 84R 84C 88C 91R 88C 87C 86C 80C 75C 90R 92R 88C 87C 87R
LOIRE VALLEY (RED)	94I 94R 95R 89T NT 89E 78E 82E NT
RHONE: COTE ROTIE/HERMITAGE	NT 96T 96E 95T 92E 97T 87E 89T 91E 92E 97T 98T 79I 89E 92E 89R 85C 96R 78C 89I 87E 95T 90E 90E 86R 90T 88C 58C 78R 92R 92R 92B 86R 84C
RHONE: CHATEAUNEUF DU PAPE	NT 95T 92E 94T 98E 93T 87E 88E 92E 88R 98E 93E 86R 98E 92R 95T 88R 90I 58C 96E 92R 90R 98R 82C 82C 90T 86C 85C 78C 65C 95R 94R 88C 60C 78C

GERMANY

MOSEL SAAR RUWER	85- 97I 100R 100T 91I 93E 95T 78I 79I NT 95R 89I 95I 82I 92R 95I 94R 92R 91I 91R 91R 76C 86E 92T 88R 91T 90R 94R 91R 87C 88R 96R 91C 92R 84C 85C
RHEINHESSEN (RIESLING)	93I 93T 92I 92R 92R 81I 83I 90T 94R 87I 93R 88R 92R 86I 92R 93T 89I 92R 95R 69C 87E 93T 87R 91T 86C 87R 88R 85C 87R 96R 90C 90R 84R 84C
GERMANY (PINOT NOIR/SPÄTBURGUNDER)	NT 95I 93I 94I 88I 90I 93E 90I 89I 90T 87T 91T 87E 88R 89R NT

2020 2019 2018 2017 2016 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2004 2003 2002 2001 2000 1999 1998 1997 1996 1995 1994 1993 1992 1991 1990 1989 1988 1987 1986

RATINGS	
96-100	Extraordinary
90-95	Outstanding
80-89	Above Average to Excellent
70-79	Average
60-69	Below Average
<59	Appalling

MATURITY	
C	Caution, may be too old
E	Early maturing and accessible
NV	Vintage not declared
I	Irregular, even among the best wines
NT	Not yet sufficiently tested to rate
R	Ready to drink
T	Still tannic, youthful, or slow to mature

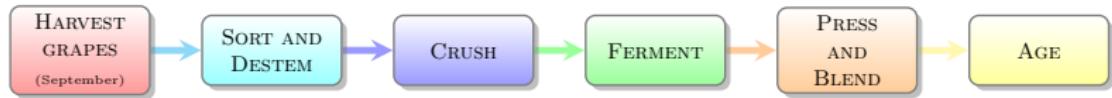
The Statistical Sommelier

- Orley Ashenfelter, Professor of Economics, Princeton University.
- Claims that he can predict quality without tasting.
- Tool used: Linear Regression
- <http://www.liquidasset.com/>

Why predict?

- ... ???
- Investment: *en primeur* or *wine futures*
 - tasting samples of wine within a year or two after it is made.
- The 1982 vintage wine of Chateau Latour that was sold at 250 pounds a case en primeur in 1983 was valued at 9000 British pounds in 2007.
- <https://www.youtube.com/watch?v=x9A0TctNAG0>

Schematic of the wine making process



Insights

Observations:

- ① Older the wine is, the greater is the value.
- ② There is still significant variation in average prices that remains unexplained.

Possible predictors:

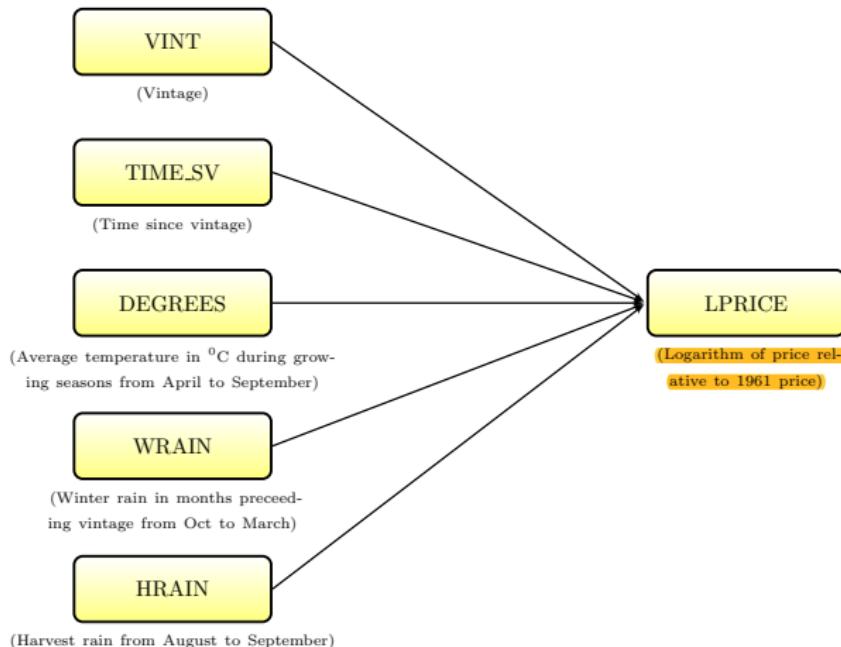
- ① Vineyard (chateau - location where wine is made)
- ② Vintage (year - time when wine is made)

The clip below gives an idea about the winemaking process in Bordeaux:

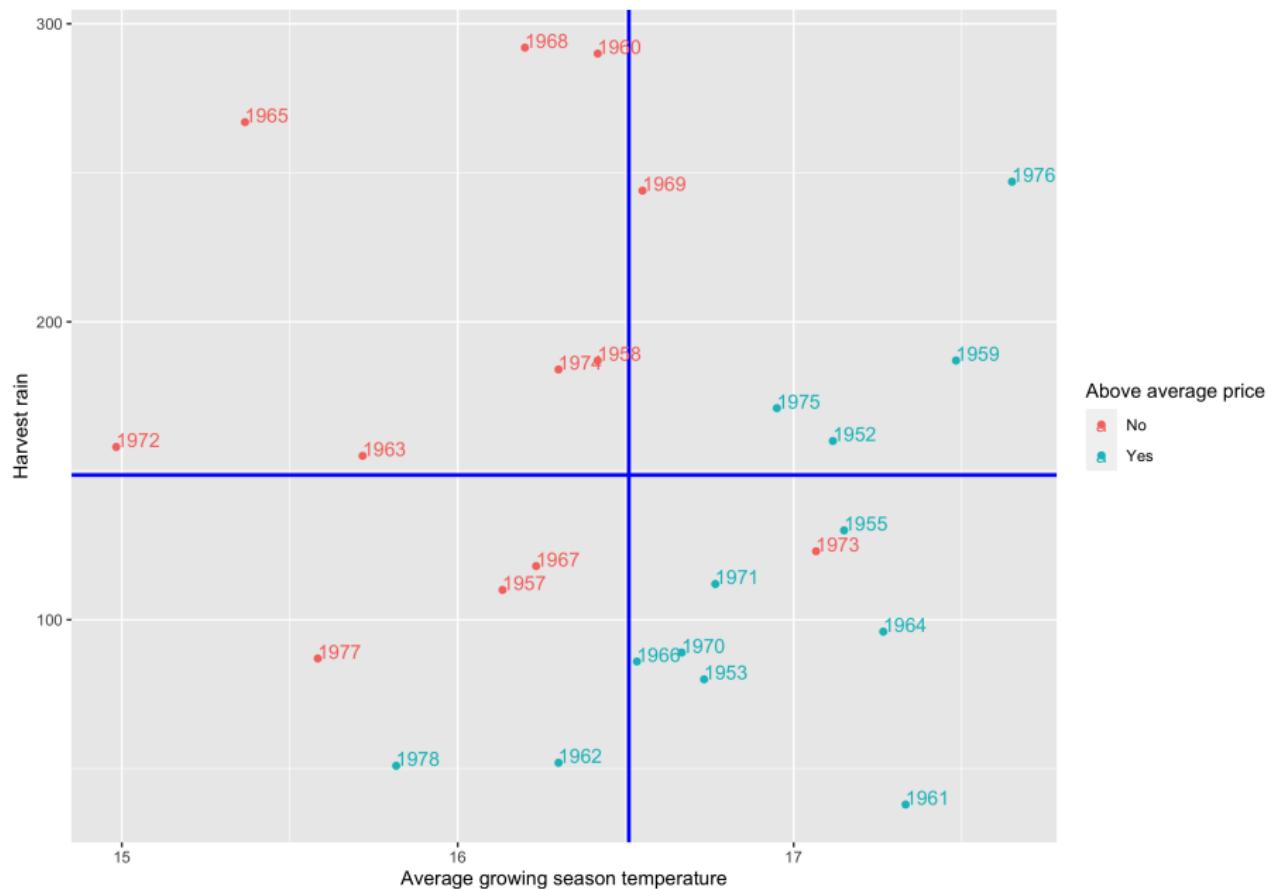
<https://youtu.be/3bI6lkIqZ9M?t=8>

Datasets and variables

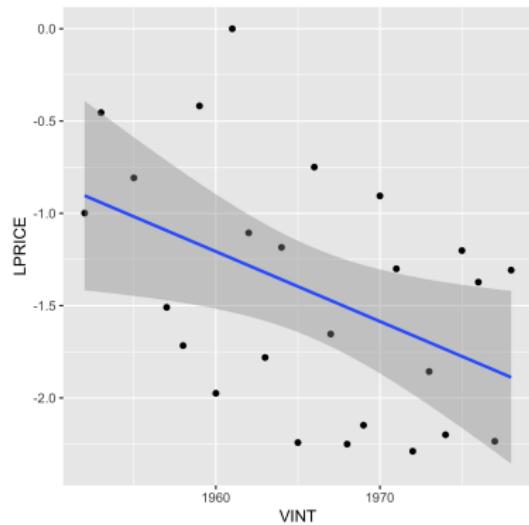
- Years: 1952 - 1989



Variables used from data in www.liquidasset.com



Linear Regression



Problem setup:

- ① n = Number of observations
- ② p = Number of predictor variables
- ③ y = Dependent variable in \mathbb{R}
- ④ x_1, \dots, x_p = Independent variables

Linear Regression

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}.$$

Linear Regression

Standard errors:

Linear Regression

Is the i^{th} variable useful in the model?

Linear Regression

Quality of fit:

Linear Regression

Quality of fit:

the best way is spend on competitive \$724.7 million, Gov. Booth Gardner, a Democrat, said in a statement, adding that allowing us to fund services that are available in need will really theme that help people." He said.

New York defends its plan to borrow \$85 million in deficit notes to cover projected shortfalls in state revenue. "It would be ludicrous to say we didn't have to do it in case the economy suddenly got worse," said Assemblyman

John C. McHugh, a Republican who is helping to draft legislation that would allow the state to borrow up to \$1 billion.

In its state liquor business to raise \$20 million. Arizona is cutting back on state spending and preparing to raise \$10 million in taxes.

Belt-tightening in some states will be accompanied by new games for managing

state budgets, but the demands

in each state were obviously

different, and the solutions

were different as well.

"It's the past two years and Congress is trying to solve its problems by cutting the budget and trying to stretch the dollars further," said Robert C. Greenlaw, managing director of the Government & International Department,

Comparing Red Bordeaux Ratings

□ Only Ashkenfield's calculation (1961 = 100)
High-Johnson's scale (scale of 0 to 10)
Robert M. Parker's 7 comment

1987	3 to 6	pleasant, soft, clean, fruity
1986	6 to 9	very good, sometimes exceptional
1985	6 to 8	solt, fragrant, very good
1984	4 to 7	austere, mediocre quality
1983	6 to 9	superior to 1981, rarely achieves greatness
1982	6 to 10	most complex and interesting wines since 1967
1981	5 to 8	lacks generosity and richness
1980	4 to 7	light and disappointing

* Ratings are for Médoc and Graves wines.

Sources: Liquid Assets Newsletter (Ashkenfield's prediction); High-Johnson's Wine Report; Parker (Elements of Wine 1987); Parker Wine Report (Oct. 1989).

Airlines Get Rules for Disabled

WASHINGTON, March 2 (AP) — The Transportation Department, recently established by President Bush, last week announced standards for air travel that are designed to accommodate disabled people on airplanes.

The announcement sought minimum requirements for the treatment of disabled passengers on commercial flights, which said the rules would prove useful to disabled people and their organizations of blind people, who will continue to face difficulties in flying those that have emergency exits.

The Transportation Department and the Federal Aviation Administration, after the standard period of public comment, will issue regulations by mid-Tuesday, would cost the airlines about \$100 million over four years on design modifications and personnel training.

What the Rules Mandate

The rules require that all airlines with more than 30 seats have accessible aircraft. They also require that aircraft that wide-body planes — generally those with 100 or more passengers — have at least one lavatory that can accommodate a handicapped passenger.

They also strictly limit airlines' ability to discriminate against disabled people, prohibit any restriction on where they sit, and require that they be seated and generally prevent airlines from discriminating against passengers due to their disabilities, according to the Transportation Department's news release.

If an airline does have such a regulation, it must be published in the *Annals of the National Institute of Health*, the journal of the National Institutes of Health.

Surgeon General Confirmed

WASHINGTON, March 3 (AP) — Dr. Louis W. Sullivan, a surgeon general at Surgeon General by the Senate on Thursday. A 45-year-old pediatrician from Chicago, Sullivan is the first female and the first Hispanic to be confirmed as surgeon general, director of the National Institute of Health.

As an orthopedic surgeon, he has performed many operations on children, especially those with birth defects. He has also been involved in research on AIDS and other diseases.

The new provision gives both manual and electric wheelchairs priority in making airplane storage areas available. It also mandates that, starting in 1993 or later will have to be built in accordance with the new rules, and older aircraft

will have to comply as they enter or are refurbished, usually about every three years.

Response From Carriers

The Air Transport Association, which represents most of the nation's carriers, and the requirements would "diminish the value of our products," said a spokesman.

According to the association, the new regulations do not change a rule of the Federal Aviation Administration that requires an estimated one-way fare to passengers who can't get to the airport by themselves.

Mark J. Hirsch, spokesman for the National Federation of the Blind, said F.A.A. rule "reflects a pattern of the F.A.A.'s rule-making," he said.

Michael S. Gitterman, spokesman for the Department of Transportation, said the agency is carrying out its responsibilities to carry on for some time.

Years to Year Variations

Most wine critics are of course at Christie's auction house in London and a leading authority on Bordeaux, which is the most important wine in the world. By taste, quality, and price, it is the most important wine in the world.

According to this "multivariate regression analysis," heavy rains in the 1980 vintage had a negative effect on wine quality, while rainfall during the growth period had a positive effect.

According to the Ashkenfield's rating, because average growing season rainfall down the 1980 Bordeaux vintage was higher than the 1989 vintage, it will be judged the wine vintage of the 1980's, and no better than

1989.

Wine Equation

The wine critic, Robert M. Parker, Jr.,

has

predicted

that

the

1989

vintage

will

be

the

best

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

the

1983

vintage

will

be

the

worst

since

1982

and

<p

The Bordeaux equation and wine critics

- ① In 1991, Ashenfelter predicted 1989 and 1990 to be exceptional years. Wine critics did not agree at that time; but there is almost unanimous agreement now that they are the best vintages of the last 50 years.
- ② Ashenfelter's model and Parker's expert opinion both agree that the 2000 and 2003 vintage would be very high quality wine.
- ③ All of this was done without tasting a single drop of the wine.