



IIT Madras
ONLINE DEGREE

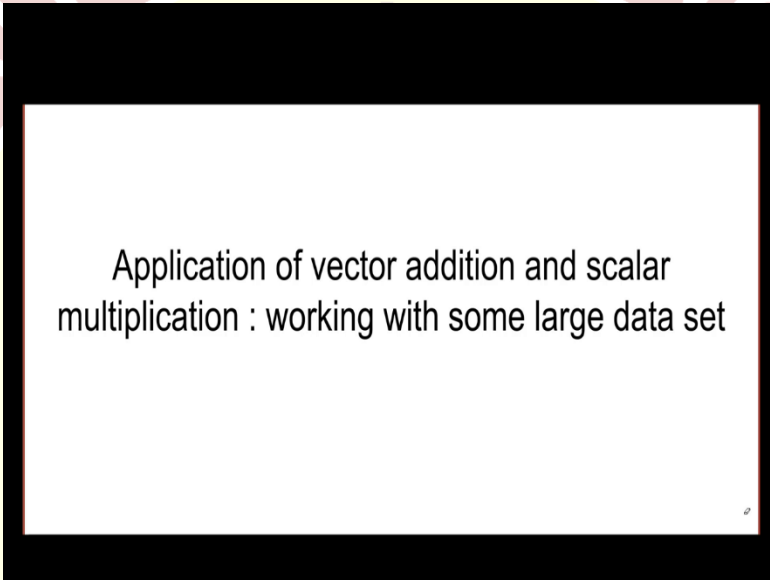
Mathematics for Data Science 2
Professor Sarang S. Sane
Department of Mathematics,
Indian Institute of Technology Madras
Mr. Subhajit Chanda
Course Instructor,

Indian Institute of Technology Madras Online Degree Program

Tutorial 05

Application of vector addition and scalar multiplication: working with some large data set

(Refer Slide Time: 00:14)



Application of vector addition and scalar
multiplication : working with some large data set

Hello. So in this video we will see an application of vector addition and scalar multiplication. So where we will consider some large dataset and we will how we can apply vector addition and scalar multiplication which you have learned in this Week 1.

(Refer Slide Time: 00:31)

Handwritten notes on the left side of the slide:

$$C_3 = \frac{1}{2}(C_1 + C_2)$$

$$\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \vdots \end{pmatrix} \begin{pmatrix} V_1 + W_1 \\ V_2 + W_2 \\ \vdots \end{pmatrix}$$

Handwritten notes on the right side of the slide:

$$\begin{matrix} C_1 & C_2 \\ V_1 & W_1 \\ V_2 & W_2 \\ V_3 & W_3 \\ \vdots & \vdots \end{matrix}$$

$$\begin{matrix} C_1 + C_2 \\ V_1 + W_1 \\ V_2 + W_2 \\ V_3 + W_3 \\ \vdots \end{matrix}$$

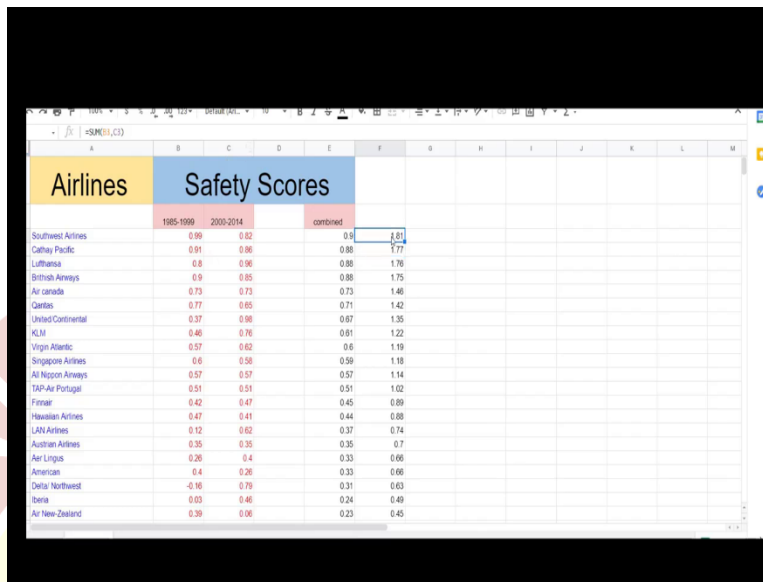
Airlines	Safety Scores		
	C ₁ 1985-1999	C ₂ 2000-2014	combined C ₃
Southwest Airlines	0.96	0.82	0.9
Cathay Pacific	0.91	0.86	0.88
Lufthansa	0.8	0.96	0.88
British Airways	0.9	0.85	0.88
Air Canada	0.73	0.73	0.73
Qantas	0.77	0.65	0.71
United/Continental	0.37	0.98	0.67
KLM	0.46	0.76	0.61
Virgin Atlantic	0.57	0.62	0.6
Singapore Airlines	0.6	0.58	0.59
All Nippon Airways	0.57	0.57	0.57
TAP-Air Portugal	0.51	0.51	0.51
Finnair	0.42	0.47	0.45
Hawaiian Airlines	0.47	0.41	0.44
LAN Airlines	0.12	0.62	0.37
Austrian Airlines	0.35	0.35	0.35
Aer Lingus	0.26	0.4	0.33
American	0.4	0.28	0.33
Delta/ Northwest	-0.16	0.79	0.31
Iberia	0.03	0.46	0.24
Air New-Zealand	0.39	0.08	0.23
Air Concorde	0	0.44	0.22
COPA	-0.05	0.49	0.22
Alaska Airlines	0.36	0	0.2
Aerolineas Argentinas	0.03	0.36	0.19
EL AL	0.01	0.37	0.19

So let us take this dataset where in the left hand side we have airlines and in the right hand side we have safety scores. These safety scores is basically an indicator whether the airline is safe or not. So it depends on the number of accidents and number of fatalities occurred in those accidents. And so where in column one we are taking time period from 1985 to 1999 and in the second column we are taking the time period from 2000 to 2014.

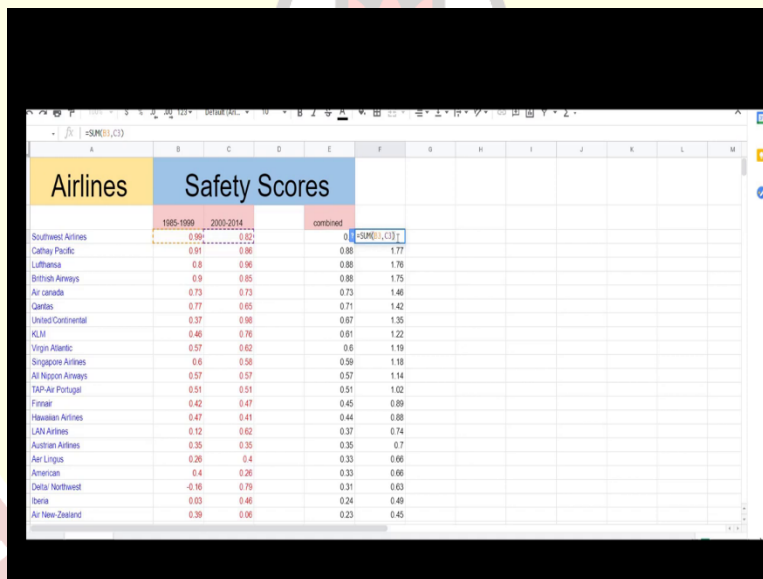
And in the third column we are seeing the combined. So let us denote this as C_3 . So in column one there are some entries V_1, V_2, V_3 and so on and in column two there are some entries like W_1, W_2, W_3 and so on. So when we are adding C_1 and C_2 , we are basically adding coordinate wise that means $V_1+W_1, V_2+W_2, V_3+W_3$ and so on, which we have seen in the lecture. And what C_3 represents, C_3 represent is the average of C_1 and C_2 .

So C_3 is nothing but half of $V_1 + W_1$. This will be in the first row, half of V_2+W_2 will be in the second row. So this is the scalar multiplication. This is the scalar. And we are multiplying with the entry in the first row which is V_1+W_1 . This is the scalar half and we are multiplying it with the vector, with the entry V_2+W_2 which is in the second row. That is how we are computing C_3 .

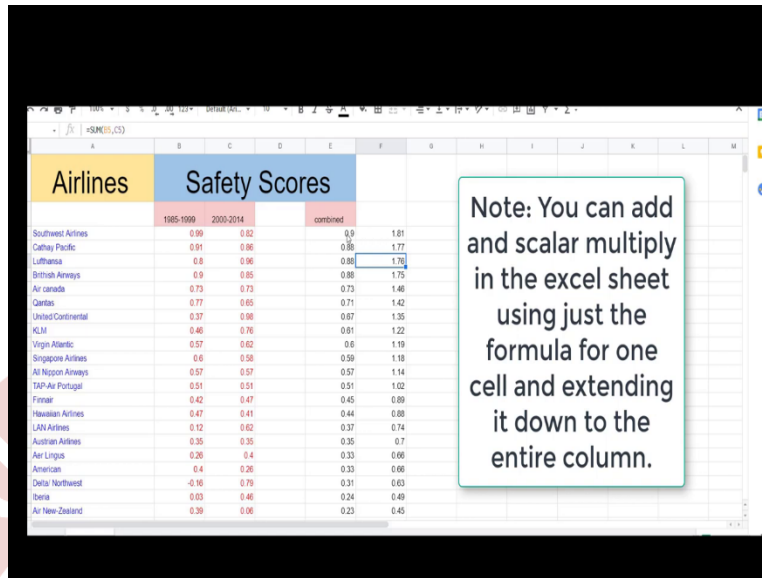
(Refer Slide Time: 02:32)



Airlines	1985-1999	2000-2014	combined
Southwest Airlines	0.99	0.82	0.9
Cathay Pacific	0.91	0.96	0.86
Lufthansa	0.8	0.96	0.88
British Airways	0.9	0.85	0.88
Air Canada	0.73	0.73	0.73
Qantas	0.77	0.65	0.71
United Continental	0.37	0.98	0.67
KLM	0.46	0.76	0.61
Virgin Atlantic	0.57	0.62	0.6
Singapore Airlines	0.6	0.58	0.59
All Nippon Airways	0.57	0.57	0.57
TAP-Air Portugal	0.51	0.51	0.51
Finnair	0.42	0.47	0.45
Hawaiian Airlines	0.47	0.41	0.44
LAN Airlines	0.12	0.62	0.37
Austrian Airlines	0.35	0.35	0.35
Aer Lingus	0.26	0.4	0.33
American	0.4	0.26	0.33
Delta/Northeast	-0.16	0.79	0.31
Iberia	0.03	0.46	0.24
Air New Zealand	0.39	0.06	0.23



Airlines	1985-1999	2000-2014	combined
Southwest Airlines	0.99	0.82	0.9
Cathay Pacific	0.91	0.96	0.86
Lufthansa	0.8	0.96	0.88
British Airways	0.9	0.85	0.88
Air Canada	0.73	0.73	0.73
Qantas	0.77	0.65	0.71
United Continental	0.37	0.98	0.67
KLM	0.46	0.76	0.61
Virgin Atlantic	0.57	0.62	0.6
Singapore Airlines	0.6	0.58	0.59
All Nippon Airways	0.57	0.57	0.57
TAP-Air Portugal	0.51	0.51	0.51
Finnair	0.42	0.47	0.45
Hawaiian Airlines	0.47	0.41	0.44
LAN Airlines	0.12	0.62	0.37
Austrian Airlines	0.35	0.35	0.35
Aer Lingus	0.26	0.4	0.33
American	0.4	0.26	0.33
Delta/Northeast	-0.16	0.79	0.31
Iberia	0.03	0.46	0.24
Air New Zealand	0.39	0.06	0.23



Airlines	1985-1999	2000-2014	combined
Southwest Airlines	0.99	0.82	0.9
Cathay Pacific	0.91	0.86	0.88
Lufthansa	0.8	0.86	0.88
British Airways	0.9	0.85	0.88
Air Canada	0.73	0.73	0.73
Qantas	0.77	0.65	0.71
United Continental	0.37	0.88	0.87
KLM	0.46	0.76	0.81
Virgin Atlantic	0.57	0.62	0.6
Singapore Airlines	0.6	0.58	0.59
All Nippon Airways	0.57	0.57	0.57
TAP Air Portugal	0.51	0.51	0.51
Finnair	0.42	0.47	0.45
Hawaiian Airlines	0.47	0.41	0.44
LIAT Airlines	0.12	0.62	0.37
Australian Airlines	0.35	0.35	0.35
Air Lingus	0.26	0.4	0.33
American	0.14	0.26	0.33
Delta/Northwest	-0.16	0.79	0.31
Iberia	0.03	0.46	0.24
Air New Zealand	0.39	0.06	0.23

Note: You can add and scalar multiply in the excel sheet using just the formula for one cell and extending it down to the entire column.

It is difficult to do it by hand. So what we are doing, so we can use this excel sheet to find the entries. So this column is denoted as B here in the excel sheet and this column is denoted as C in the excel sheet. So what we have done here, you can see that we are taking B_3 , sum of B_3 and C_3 . B_3 is this one because this is the third row.

First two row are used for this naming of this column. So we are starting from the third row. So we are taking sum of B_3 and C_3 and we write it here and similarly this will be sum of B_4 and C_4 , this will be sum of B_5 and C_5 and so on. And then if we just scalar, the if we just multiply these entries with half, which is a scalar in real number basically, then we will get this column, the combined column.

So that is how you can see that using this vector addition which eventually we will use in programming then it will be easier to work with this large dataset or even you can use this excel sheet what I am using here to calculate the sum of two vectors, large vectors basically or calculating the scalar multiplication of those large vector. Thank you.