Welcome

Your Lecturer for this session

Sanana Mwanawina

Lecture - Housekeeping

- ❏ The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly.
- ❏ No question is daft or silly - **ask them!**
- ❏ There are Q/A sessions midway and at the end of the session, should you wish to ask any follow-up questions.
- ❏ You can also submit questions here: hyperiondev.com/sbc4-ds-questions
- ❏ For all non-academic questions, please submit a query: hyperiondev.com/support
- ❏ Report a safeguarding incident: hyperiondev.com/safeguardreporting
- ❏ We would love your feedback on lectures: https://hyperionde.wufoo.com/forms/zsgv4m40ui4i0g/

Objectives

- Learn about a simple machine learning algorithm, the regression analysis
- Understand how it is used to estimate the relationship between variables.

Regression Analysis

★ Regression analysis is a statistical process used to estimate relationships between variables.

★ There are two types of linear regression: simple linear regression, which is the focus of this task, and multiple linear regression, which will be taught in the next task. We will teach these concepts using examples involving limited numbers of variables.
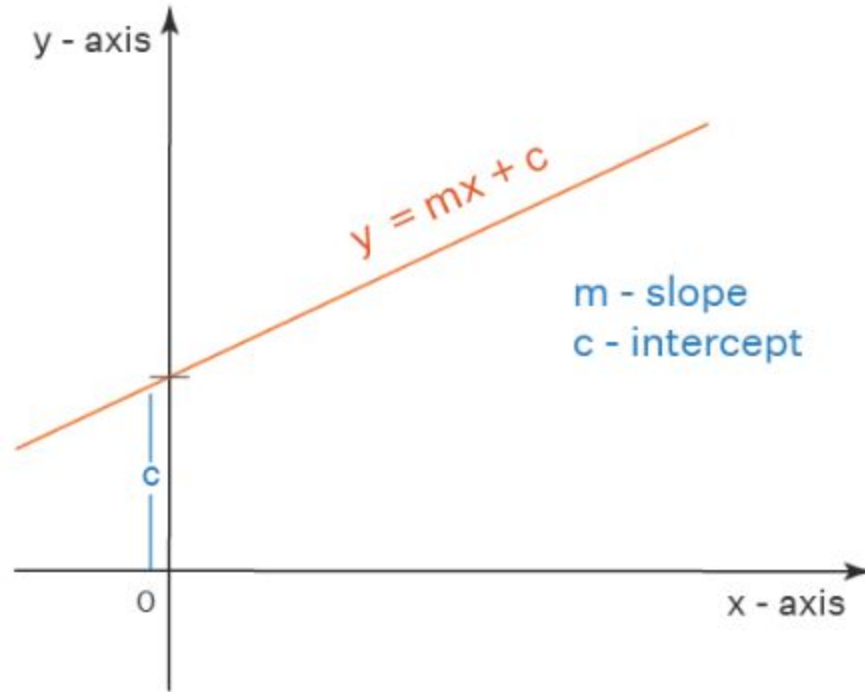
Linear Regression

★ Suppose that we have been asked by a client to give advice on how to advertise their product most effectively. The client can offer us some data on which to base our recommendation. She offers us the sales and advertising budgets of the product in 200 markets, where the advertising budget is for television ads.

	1	TV	Sales
1		TV	Sales
2	1	230.1	22.1
3	2	44.5	10.4
4	3	17.2	9.3
5	4	151.5	18.5
6	5	180.8	12.9
7	6	8.7	7.2
8	7	57.5	11.8
9	8	120.2	13.2
10	9	8.6	4.8
11	10	199.8	10.6
12	11	66.1	8.6
13	12	214.7	17.4
14	13	23.8	9.2
15	14	97.5	9.7
16	15	204.1	19
17	16	195.4	22.4
18	17	67.8	12.5

Linear Regression

★ Common sense suggests that spending more money on advertising will increase sales and that different kinds of advertising do so at various rates. Indeed, as seen in the graphs below, the data show that the higher the budget, the higher the sales.

Linear Regression

Linear Regression

★ We start by expressing our assumption of a relationship between sales and advertising in the following general mathematical form: $Y = f(X)$ Here Y are the sales, X is the TV budget.

★ $f(x) = y = mx + c$

$f(x)$: Sales $= \beta_1$ TVbudget $+ \beta_0$

★ The purpose of simple linear regression is to find the straight line that "best fits" the data.

Linear Regression

★ Scenario:

Management at a firm you work at want to invest $100 000 towards the TV advertising budget. They ask you to give them an estimate of how many sales this proposed budget will bring.

Hyperiondev

Q & A Section

Please use this time to ask any questions relating to the topic explained, should you have any

Hyperiondev

Thank you
for joining us