Prediction

Prof. Abby Stylianou

Logistics

- Homework 2 due Thursday before class

`group` & `pivot` for cross-classification

- Classification: assign individuals to different groups based on shared properties

`group` & `pivot` for cross-classification

- Classification: assign individuals to different groups based on shared properties

- When individuals have multiple features, there are many different ways to classify them.

`group` & `pivot` for cross-classification

- Classification: assign individuals to different groups based on shared properties

- When individuals have multiple features, there are many different ways to classify them.
  - e.g., we have a population of college w/ a major and the number of years in college
  - students could be classified by major, or by year, or by a combination of major and year

`group` & `pivot` for cross-classification

- Classification: assign individuals to different groups based on shared properties

- When individuals have multiple features, there are many different ways to classify them.
  - e.g., we have a population of college w/ a major and the number of years in college
  - students could be classified by major, or by year, or by a combination of major and year

- `group` and `pivot` are table operations that allow us to classify individuals according to multiple variable (or to 'cross-classify' them)

`group` & `pivot` for cross-classification

Quick demo to refresh our memory of group and pivot

Joining Tables by Columns

- When you have related data in multiple tables, you can ‘join’ by shared column

```
Table_1.join(‘Table_1_column_label’, Table_2,
‘Table_2_column_label')
```

“Data Mining” and Prediction

- “Data Mining” attempts to extract patterns from data

“Data Mining” and Prediction

- “Data Mining” attempts to extract patterns from data
    - Associative patterns

        What data attributes occur together ?

“Data Mining” and Prediction

- “Data Mining” attempts to extract patterns from data
    - Associative patterns

      What data attributes occur together ?
    - Classification

      What indicates a given category ?

“Data Mining” and Prediction

- “Data Mining” attempts to extract patterns from data
  - Associative patterns

    What data attributes occur together ?
  - Classification

    What indicates a given category ?
  - Sequential/temporal patterns

    What sequences of events occur frequently ?

Example Patterns

- Associative pattern

  When Bob is in the living room he likes to watch TV and eat popcorn with the light turned off.

Example Patterns

- Associative pattern

  When Bob is in the living room he likes to watch TV and eat popcorn with the light turned off.

- Classification

  Action movie fans like to watch Terminator, drink beer, and have pizza.

Example Patterns

- Associative pattern

  When Bob is in the living room he likes to watch TV and eat popcorn with the light turned off.

- Classification

  Action movie fans like to watch Terminator, drink beer, and have pizza.

- Sequential patterns

  After coming out of the bedroom in the morning, Bob turns off the bedroom lights, then goes to the kitchen where he makes coffee, and then leaves the house.

Data Mining and Prediction

- Prediction attempts to form patterns that help to predict the next event(s) given the available input data.

Data Mining and Prediction

- Prediction attempts to form patterns that help to predict the next event(s) given the available input data.
  - Deterministic predictions

    If Bob leaves the bedroom before 7:00 am on a workday, then he will make coffee in the kitchen.

Data Mining and Prediction

- Prediction attempts to form patterns that help to predict the next event(s) given the available input data.
  - Deterministic predictions

    If Bob leaves the bedroom before 7:00 am on a workday, then he will make coffee in the kitchen.
  - Probabilistic sequence models

    If Bob turns on the TV in the evening then he will 80% of the time go to the kitchen to make popcorn.

What to Predict

- Behavior of Individuals
  - Location
  - Tasks / goals
  - Actions
- Behavior of the Environment
  - Device behavior (e.g. heating, AC)
  - Interactions

Example: Location Prediction

- Where will Bob go next?
- $\text{Location}_{t+1} = f(x)$
- Input data x:
  - Location_t, Location_{t-1}, …
  - Time, date, day of the week
  - Sensor data

Example: Location Prediction

Time	Date	Day	$Location_t$	$Location_{t+1}$
6:30	02/25	Monday	Bedroom	Bathroom
7:00	02/25	Monday	Bathroom	Kitchen
7:30	02/25	Monday	Kitchen	Garage
17:30	02/25	Monday	Garage	Kitchen
18:00	02/25	Monday	Kitchen	Bedroom
18:10	02/25	Monday	Bedroom	Living room
22:00	02/25	Monday	Living room	Bathroom
22:10	02/25	Monday	Bathroom	Bedroom
6:30	02/26	Tuesday	Bedroom	Bathroom

Example: Location Prediction

- Learned pattern
  - If Day = Monday…Friday

    & Time > 0600

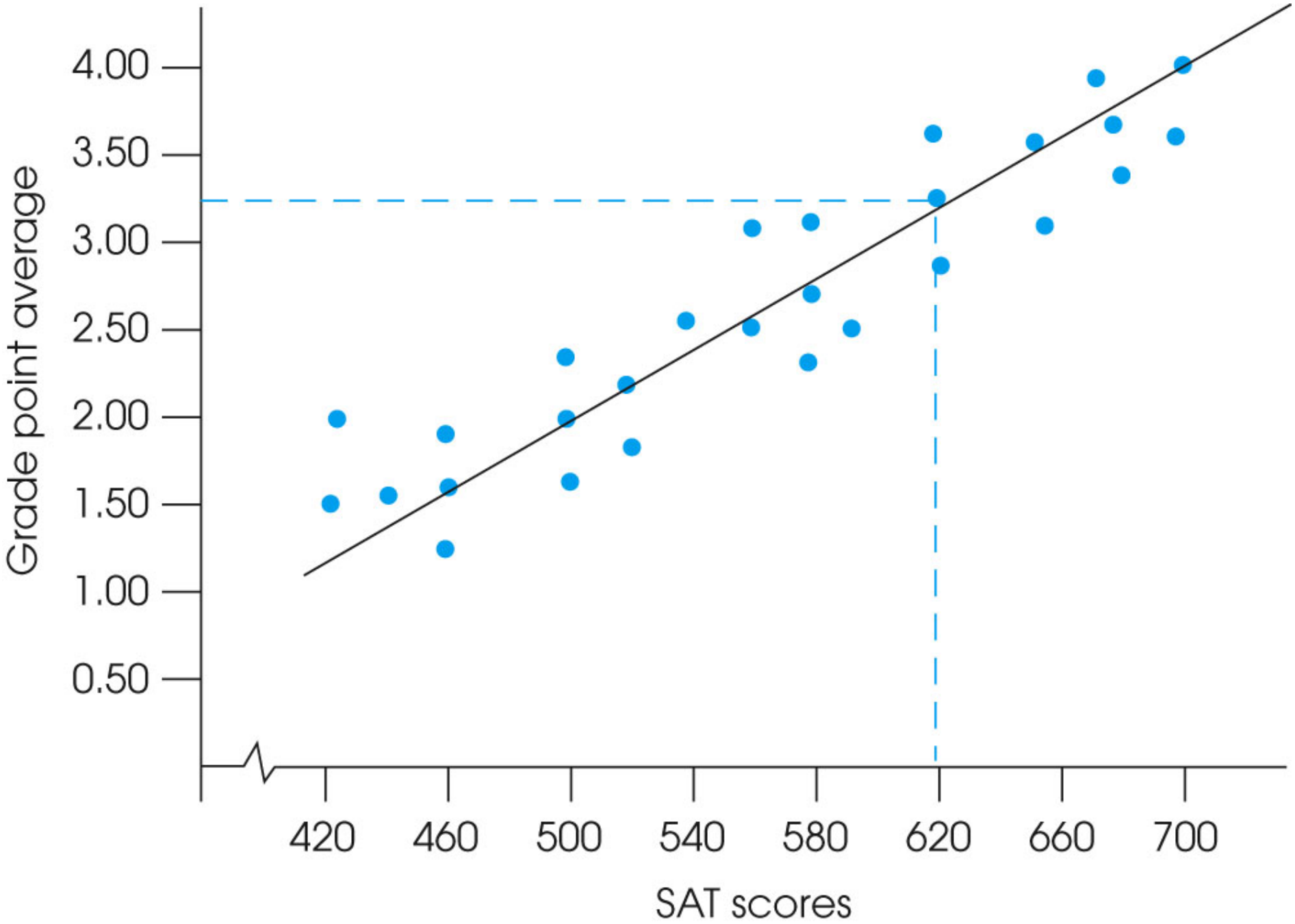
    & Time < 0700

    & Location_t = Bedroom

    Then Location_{t+1} = Bathroom

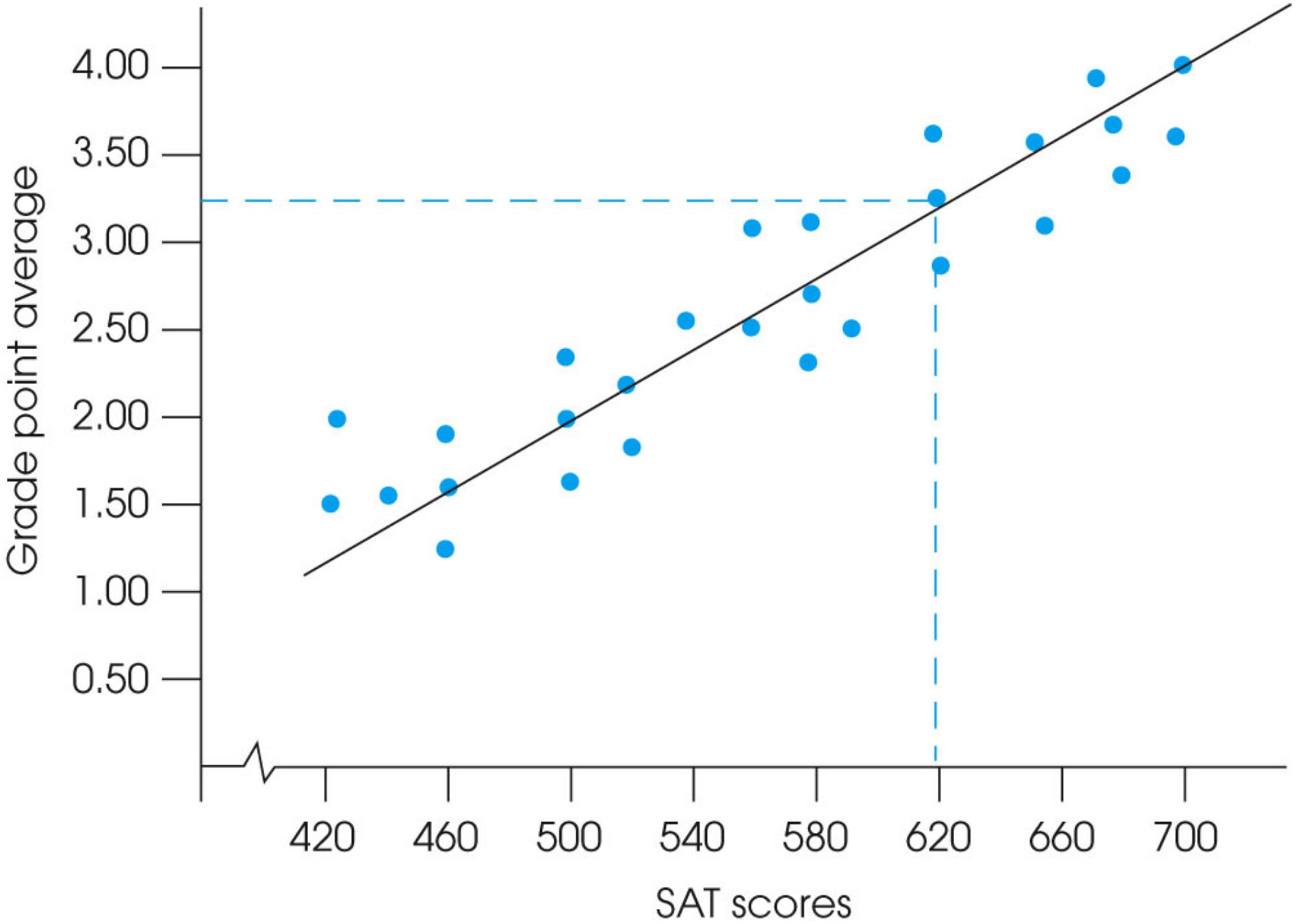
Prediction Techniques

- Classification-Based Approaches
  - Nearest Neighbor
  - Neural Networks
  - Bayesian Classifiers
  - Decision Trees
- Sequential Behavior Modeling
  - Hidden Markov Models
  - Temporal Belief Networks


Grade point average
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
420
460
500
540
580
620
660
700
SAT scores

Introduction to Linear Regression

- Any straight line can be represented by an equation of the form $y = mx + a$, where m and a are constants.
- b: slope constant, determines the direction and degree to which the line is tilted ($= \Delta y - \Delta x$)
- a: the Y-intercept, determines the point where the line crosses the Y-axis.

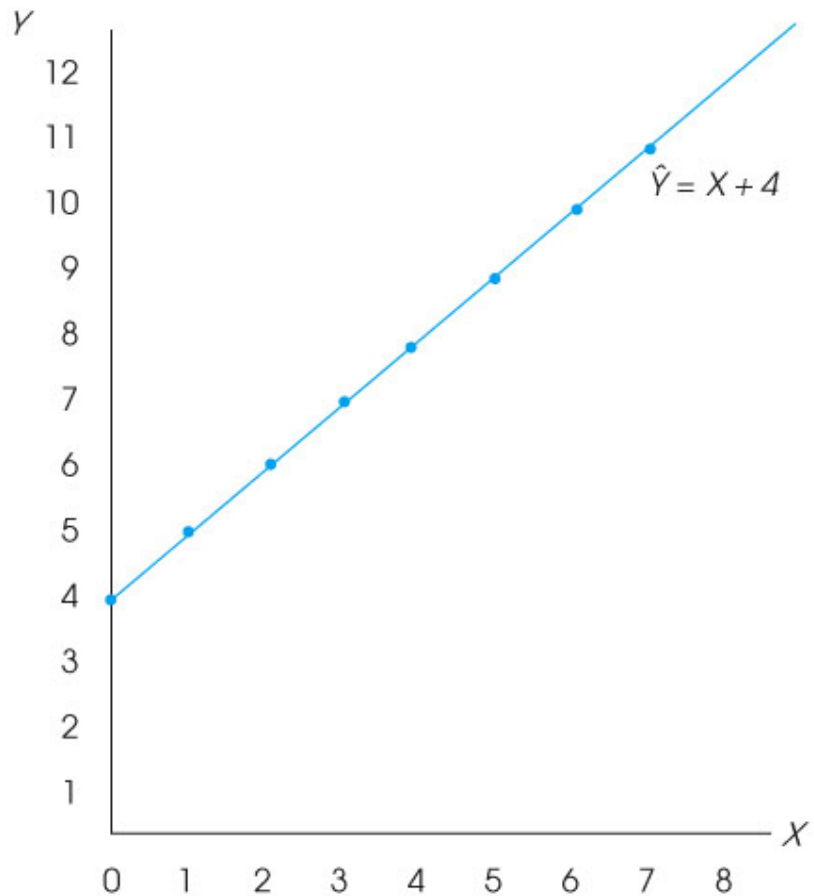

Grade point average
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
420
460
500
540
580
620
660
700
SAT scores

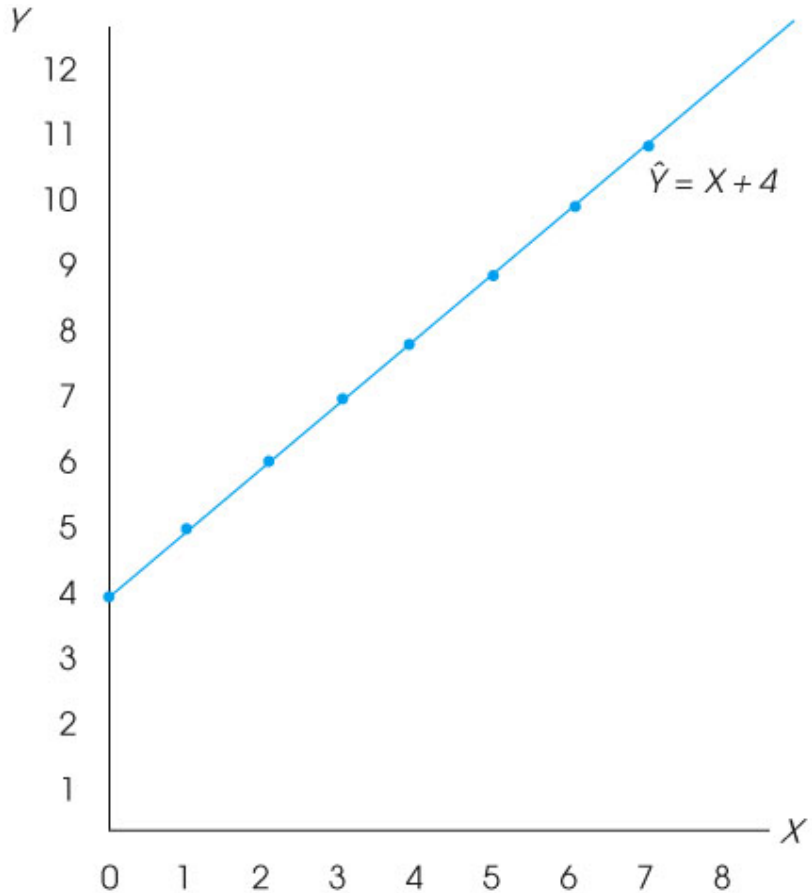
Introduction to Linear Regression

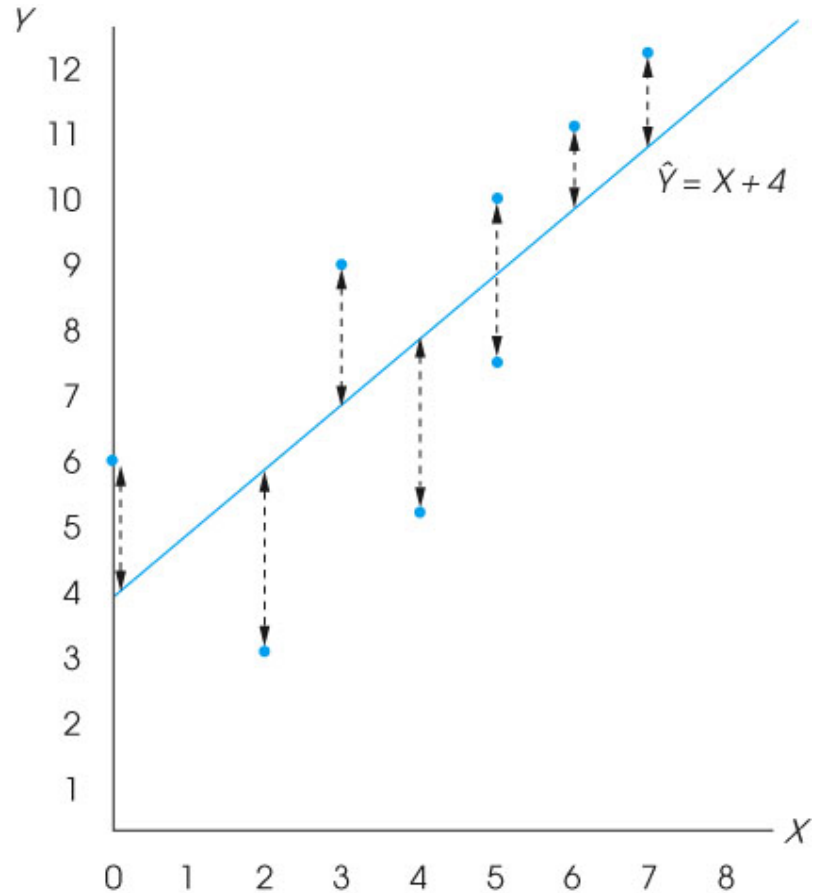
- How well a set of data points fits a straight line can be measured by calculating the distance between the data points and the line.

- The total error between the data points and the line is obtained by squaring each distance and then summing the squared values.

- Simple linear regression: find the equation of the straight line that produces the minimum sum of squared errors.


Y
12
11
10
9
8
7
6
5
4
3
2
1
$\hat{Y} = X + 4$
X
0
1
2
3
4
5
6
7
8


Y
12
11
10
9
8
7
6
5
4
3
2
1
$\hat{Y} = X + 4$
X
0 1 2 3 4 5 6 7 8


Y
12
11
10
9
8
7
6
5
4
3
2
1
$\hat{Y} = X + 4$
X
0 1 2 3 4 5 6 7 8

Introduction to Linear Regression

- Can use this ‘line’ to make predictions about the values of unseen or new data

Prediction: Sir Francis Galton

- 1822 – 1911 (knighted in 1909)
- A pioneer in making predictions
- Particular interest in heredity (father of eugenics)
- Charles Darwin's half-cousin