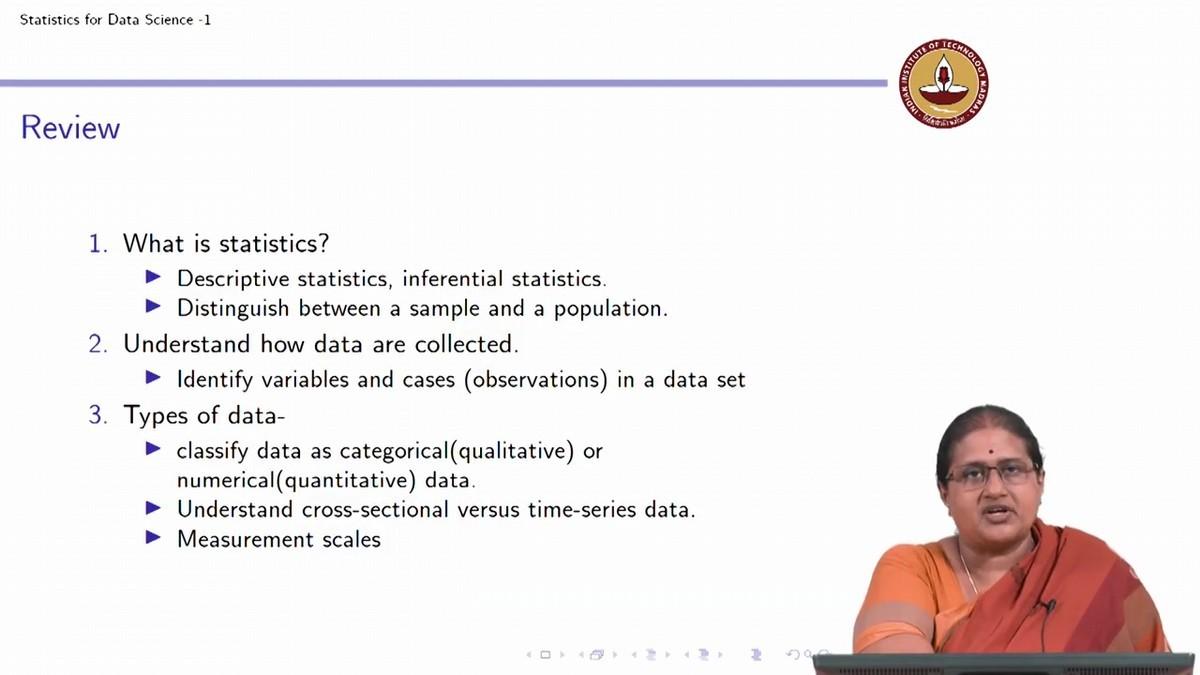
**Statistics for Data Science - 1 Prof. Usha Mohan**

**Department of Management Studies Indian Institute of Technology, Madras Lecture No. 2.1**

**Describing Categorical Data – Frequency distributions**

पिछले लेक्चर (lecture) में, आपको मूल रूप से स्टेटिस्टिक्स (statistics) की दो महत्वपूर्ण शाखाओं से परिचित कराया गया था, जो वर्णनात्मक स्टेटिस्टिक्स (statistics) और अनुमानात्मक स्टेटिस्टिक्स (statistics) हैं। आप यह भी समझ गए हैं कि सैंपल (sample) और पापुलेशन (population) में क्या अंतर है।

(Refer Slide Time: 00:32)

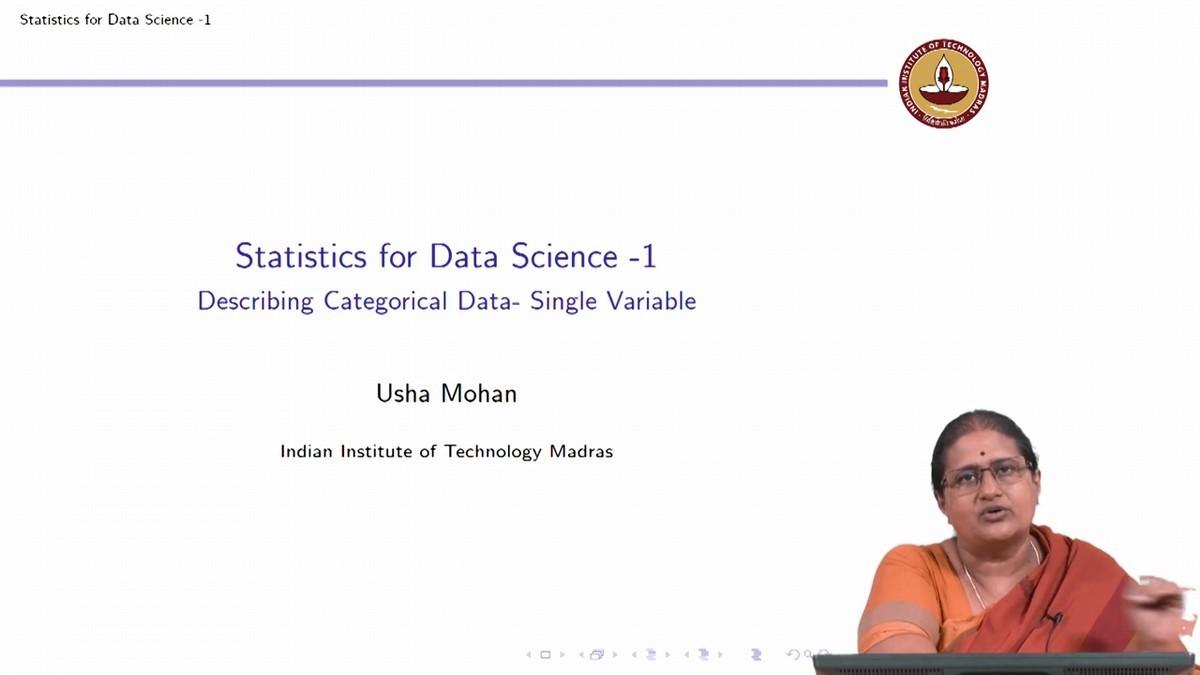


लेकिन हमने अपने परिचय को इस विशेष पाठ्यक्रम के लिए आवश्यक सैंपल (sample) और पापुलेशन (population) तक सीमित रखा है। बेशक, उन्नत पाठ्यक्रमों में आप सैंपल (sample) और पापुलेशन (population) के बारे में अधिक जानेंगे। फिर हम डेटा को पहचानने और समझने के लिए आगे बढ़े। इस पाठ्यक्रम के इस उद्देश्य के लिए फिर से हम खुद को केवल स्ट्रक्चर्ड डेटा (structured data) तक सीमित कर रहे हैं।

स्ट्रक्चर्ड डेटा (structured data) में, मेरे पास एक तालिका के रूप में डेटा (data) है जहां वेरिएबल (variable) एक कॉलम (column) में दर्ज किए जाते हैं और अवलोकन ऐसे मामले होते हैं जो डेटा (data) सेट में पंक्तियों द्वारा दर्शाए जाते हैं। हमने डेटा (data) के प्रकारों को देखा, मोटे तौर पर आप डेटा (data) को श्रेणीबद्ध डेटा (data) और संख्यात्मक डेटा (data) के रूप में वर्गीकृत करते हैं। हमने यह भी समझा कि क्रॉस-सेक्शनल डेटा (cross-sectional data) और टाइम सीरीज़ डेटा (time-series data) में क्या अंतर है।

और अंत में, आपको इस समय तक पता चल जाना चाहिए कि माप के पैमाने क्या हैं, मेरा मतलब है कि जब आपके पास श्रेणीबद्ध डेटा होता है, तो आपको पता होना चाहिए कि यह एक नॉमिनल वेरिएबल (nominal variable) या एक ऑर्डिनल वेरिएबल (ordinal data) है और जब संख्यात्मक की बात आती है, तो आपको पता होना चाहिए कि इसमें अंतर कैसे करना है। आपको पता होना चाहिए कि यह अंतराल वेरिएबल (variable) है या अनुपात वेरिएबल (variable)। इस समय आपको यही जानना चाहिए।

(Refer Slide Time: 01:50)



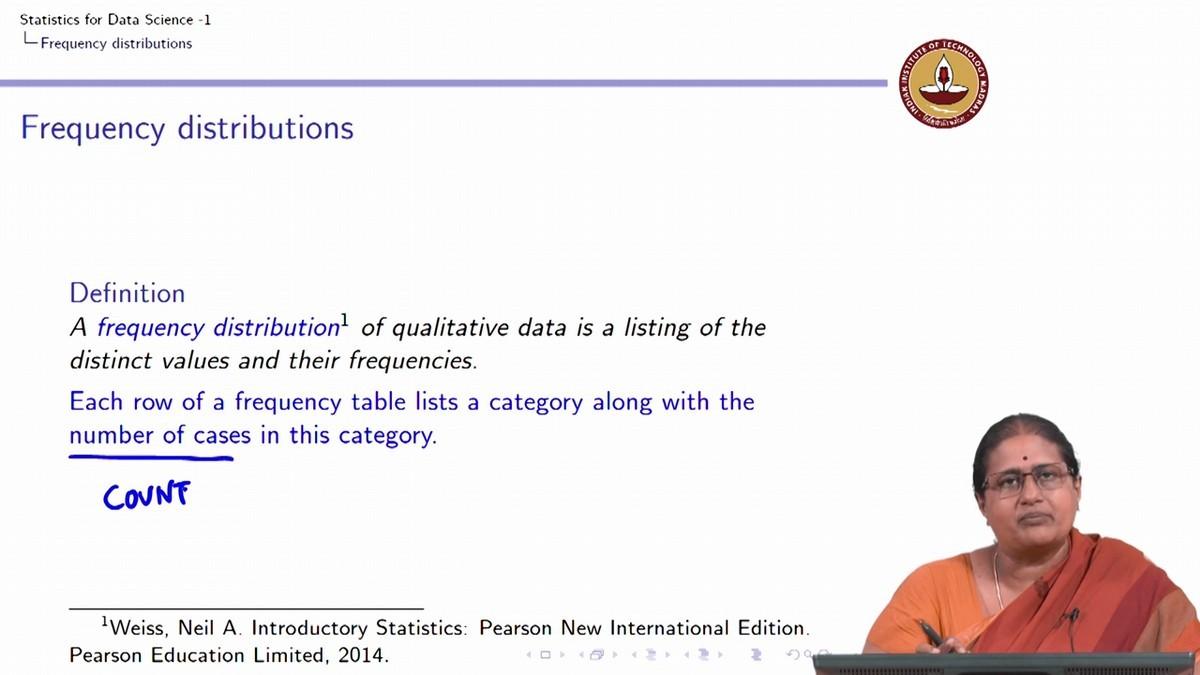
आगे बढ़ते हुए, आप समझेंगे कि श्रेणीबद्ध डेटा (data) का वर्णन कैसे किया जाता है। अब, इस मॉड्यूल (module) में आप पहले समझने जा रहे हैं, हम एक वेरिएबल (variable) के लिए श्रेणीबद्ध डेटा (data) का वर्णन करने के साथ शुरू करते हैं, और फिर हम एसोसिएशन (association) के उपायों को देखेंगे जब उनके पास एक से अधिक वेरिएबल (variable) होंगे। तो, आज हम जो करने जा रहे हैं, हम यह समझने जा रहे हैं कि श्रेणीबद्ध डेटा (data) का वर्णन कैसे किया जाता है।

(Refer Slide Time: 2:13)



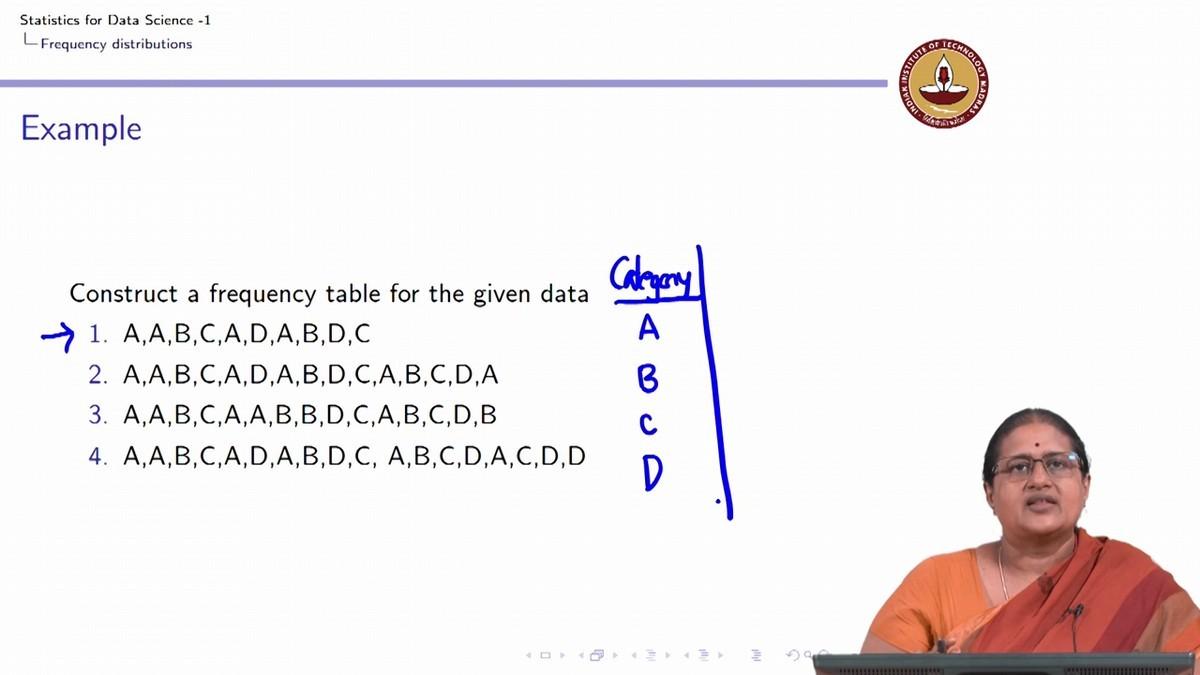
इसलिए, हम वही शुरू करते हैं जिसे हम फ्रीक्वेंसी (frequency) वितरण के रूप में समझते हैं।

(Refer Slide Time: 02:20)



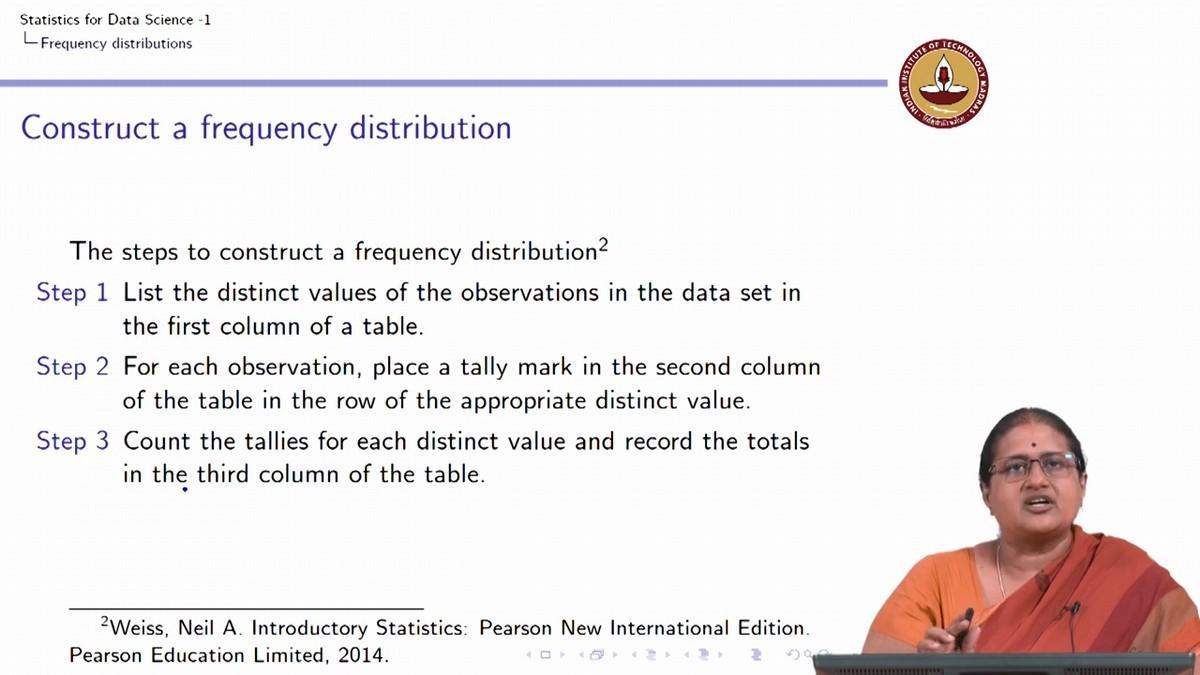
अब, फ्रीक्वेंसी (frequency) की परिभाषा, परिभाषा अलग-अलग मानों और उनकी आवृत्तियों की एक सूची है। फ्रीक्वेंसी (frequency) से हमारा क्या तात्पर्य है? फ्रीक्वेंसी (frequency) कुछ और नहीं बल्कि गिनती है, गिनती के अलावा कुछ नहीं। और अलग-अलग मानों से, आपका मतलब है कि श्रेणीबद्ध वेरिएबल (variable) वास्तव में कौन से विशिष्ट मान लेता है। इससे हमारा यही तात्पर्य है। अब, फ्रीक्वेंसी (frequency) तालिका की प्रत्येक पंक्ति, मामलों की संख्या या मामलों की संख्या के साथ एक श्रेणी सूचीबद्ध करती है। जिस क्षण मैं कहती हूं कि जितने मामले हैं, मैं सिर्फ यह कहना चाहती हूं, कि उस विशेष श्रेणी में कितने मामले हैं।

(Refer Slide Time: 03:14)



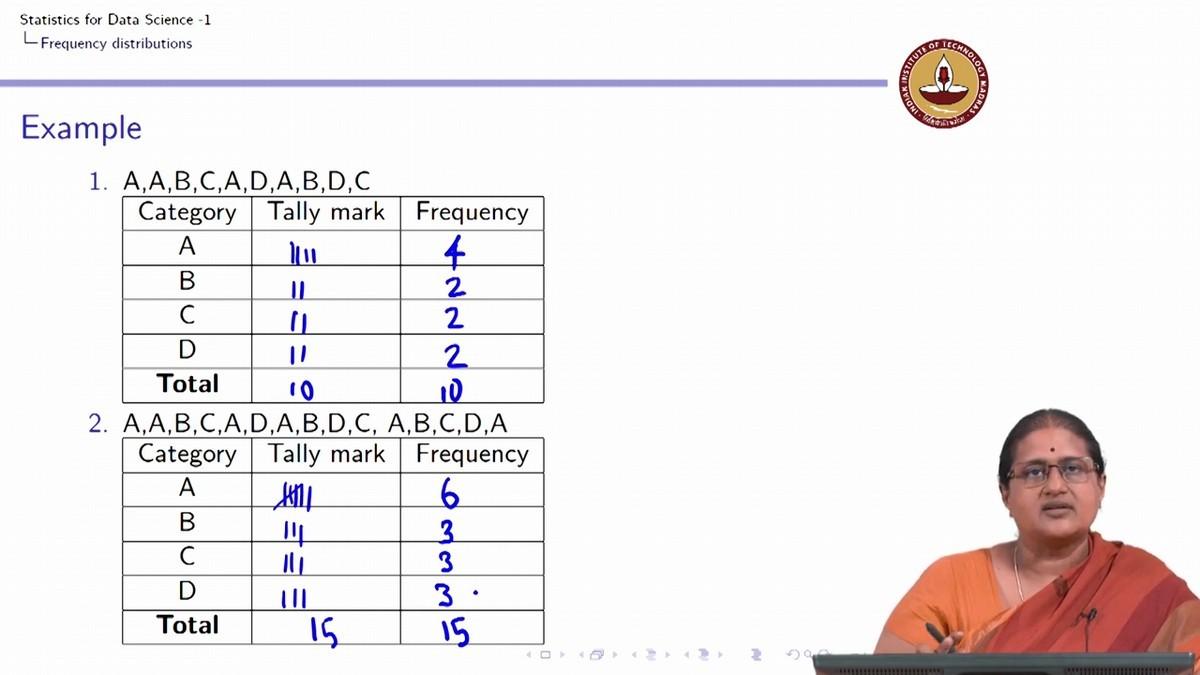
इसलिए, उदाहरण के लिए, आइए इस सरल दिए गए डेटा के लिए फ्रीक्वेंसी (frequency) तालिका का निर्माण देखें। श्रेणी सिर्फ अक्षर है। मेरे पास अक्षर हैं, मेरे पास 4 श्रेणियां हैं, मैं उन्हें A, B, C और D कह सकती हूं।

(Refer Slide Time: 3:34)



तो, मैं एक फ्रीक्वेंसी (frequency) वितरण कैसे बनाऊं? प्रेक्षण के विशिष्ट मूल्यों की सूची बनाइए। यहाँ, मेरे अवलोकन के विशिष्ट मूल्य क्या हैं? पहले उदाहरण में, मेरे अवलोकन के विशिष्ट मूल्य A, B, C और D हैं। मैं उन्हें सूचीबद्ध करती हूं और मैं इसे एक श्रेणी के रूप में लिखती हूं। यह पहली चीज है जो मैं करती हूं। वह मेरा पहला कदम है। पहला कदम विशिष्ट मूल्यों को सूचीबद्ध करना है और यह मेरा पहला कॉलम है।

(Refer Slide Time: 4:06)



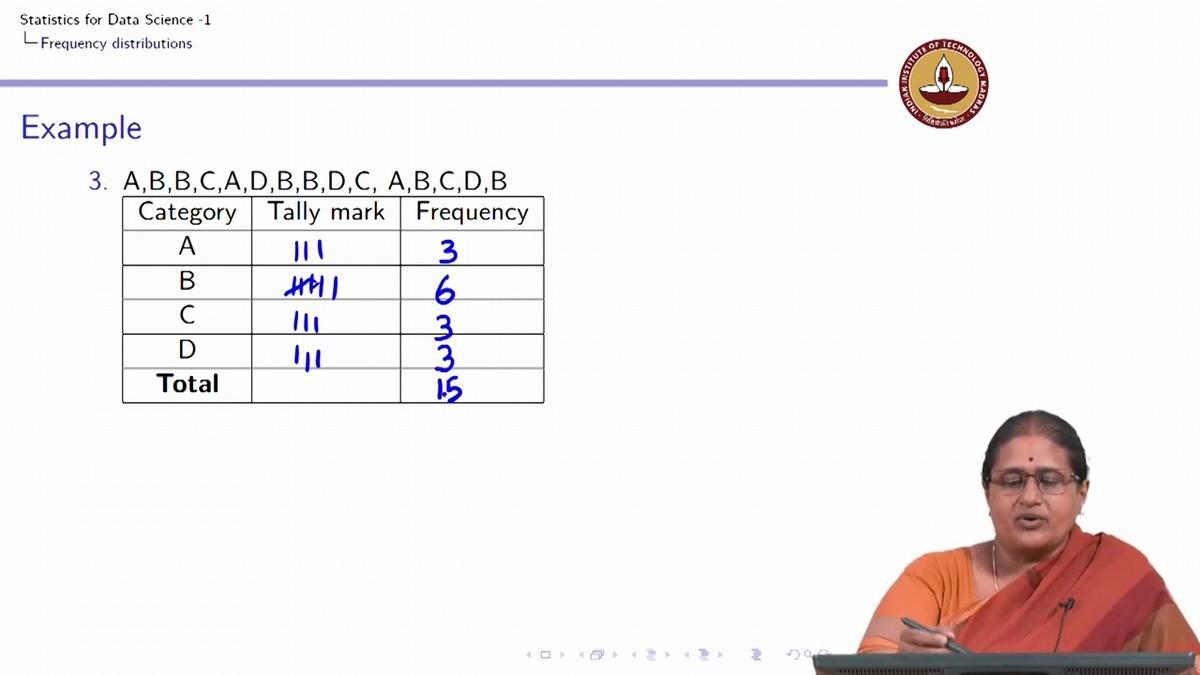
तो आप देख सकते हैं कि पहले कॉलम (column) में, मैंने अलग-अलग मान सूचीबद्ध किए हैं जो A, B, C और D हैं। प्रत्येक अवलोकन के लिए, दूसरे कॉलम में एक मिलान चिह्न लगाएं। तो A के लिए, मैं एक मिलान चिह्न लगाती हूं, जिसे मैं यहां चिह्नित कर रही हूं, दूसरा अवलोकन फिर से A है, मेरे पास फिर से एक मिलान चिह्न है, तीसरा अवलोकन B है, चौथा अवलोकन C है, पांचवां अवलोकन फिर से A है, छठा अवलोकन है D, सातवां अवलोकन A है, फिर मेरे पास B है, फिर मेरे पास D है, मेरे पास C है।

तो, ये टैली मार्क्स (tally marks) हैं जो मेरे पास होने वाले हैं। तो, वह चरण दो है। प्रत्येक प्रेक्षण के लिए एक मिलान चिह्न लगाएं और फिर टैलीस (tallies) की संख्या गिनें। इसलिए, यदि मैं टैलीस (tallies) की संख्या गिनती हूं, तो फ्रीक्वेंसी (frequency), या A की संख्या 4 है, B की संख्या 2 है, C की संख्या 2 है, और D की संख्या 2 है। यही वह है जिसे मैं फ्रीक्वेंसी (frequency) वितरण के रूप में संदर्भित करती हूं , जहां कॉलम (column) 1 में अलग-अलग मान दिए गए हैं, कॉलम (column) 2 में मिलान चिह्न और कॉलम (column) 3 में गिनती है।

आइए अब इस उदाहरण को देखें। यदि आप इस उदाहरण को देखें, तो फिर से मेरे पास एक मिलान चिह्न है। मैं इसे चुपचाप, इस समय तेजी से करुँगी, A, A, B, C, A, D, A, B, D, C। अब यह A, मैं इसे पार करती हूं, क्योंकि यह मेरा पांचवां मूल्य है। जब भी मेरे पास 4 से अधिक होते हैं, तो मेरे पास एक टैली (tally) मार्क होता है। मैं इसे पांचवें मान के रूप में पार कर रही हूं, फिर मेरे पास B, C, D, और A है। तो, यह 5 + 1 मुझे 6 का मान देता है, यह 3 है, यह 3 है, यह 3 है .

तो, मेरे पास इस मामले में कुल 15 अवलोकन हैं, जहां यह वितरण है, श्रेणी A छह बार होती है, B 3 बार होती है, C तीन बार होती है, और D तीन बार होती है। इसलिए, यदि आप इस डेटा (data) को देखें, जहां A का मिलान चिह्न 4 था, B का 2 था, C का 2 था, और D के पास कुल 10 था। यह 4, 2, 2, 2 था। यह है, 6, 3, 3, 3.

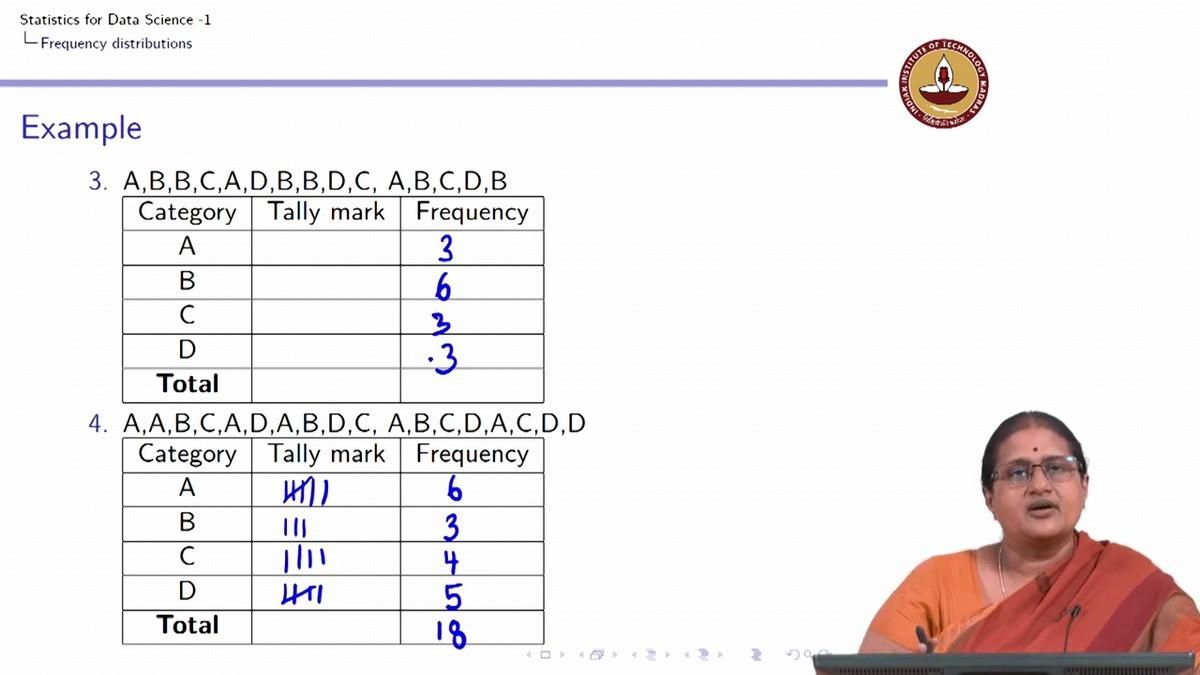
(Refer Slide Time: 06:40)



अब, आइए इसे यहां फिर से देखें। 15 अंक का वही डेटा (data), लेकिन अब मेरे पास A है, मेरे पास B, B है, मेरे पास C है, मेरे पास A है, मेरे पास D है। मेरे पास B, B, D, C, एक A है, एक B, एक C, एक D, एक B। तो, इसमें A 3 बार दिखाई दे रहा है, B 6 बार दिखाई दे रहा है, C 3 बार दिखाई दे रहा है, और D 3 बार दिखाई दे रहा है। मेरे पास फिर से कुल 15 अवलोकन हैं। इसलिए, यदि आप इस उदाहरण की तुलना पिछले उदाहरण से करते हैं, तो आप देखते हैं कि उदाहरण 2 और उदाहरण 3 के बीच एकमात्र अंतर यह है कि A यहां 6 बार और B 3 बार प्रकट होता है, C और D प्रत्येक में 3 बार दिखाई देता है।

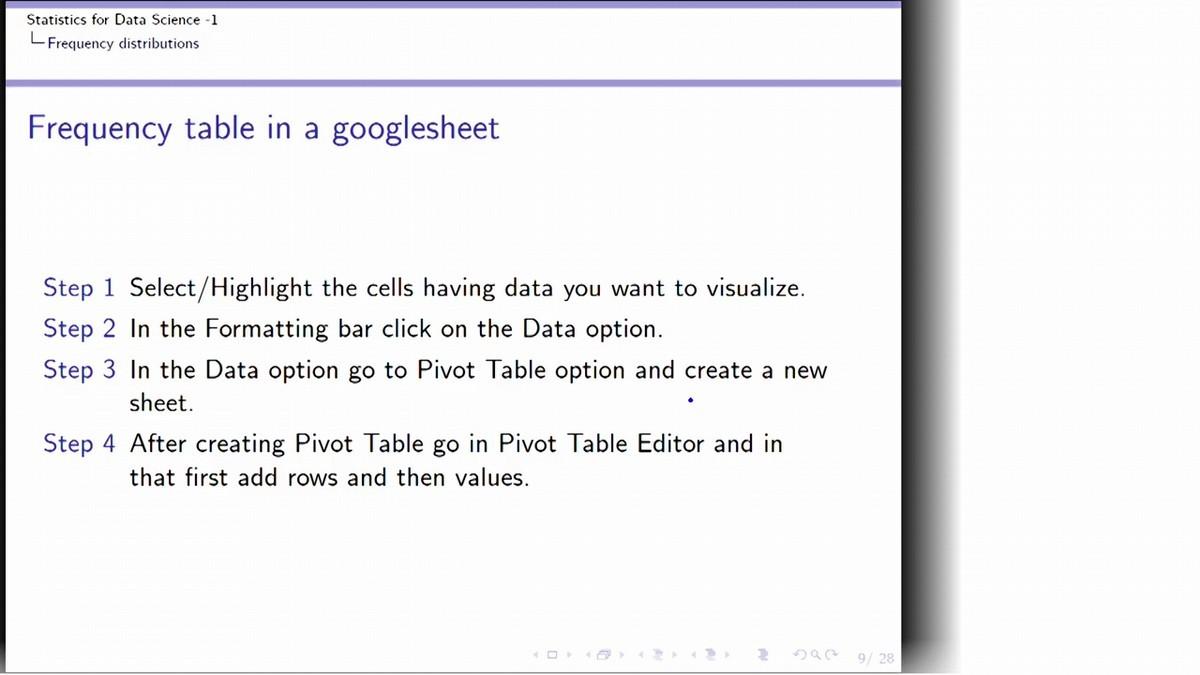
इस उदाहरण में, यह A और B के बीच फ़्लिप (flip) हो गया है जिसमें A 3 बार दिखाई दे रहा है, B 6 बार दिखाई दे रहा है, और C और D प्रत्येक 3 बार दिखाई दे रहा है। उन दोनों के 15 अवलोकन हैं, और यही दिया गया है।

(Refer Slide Time: 7**:**48)



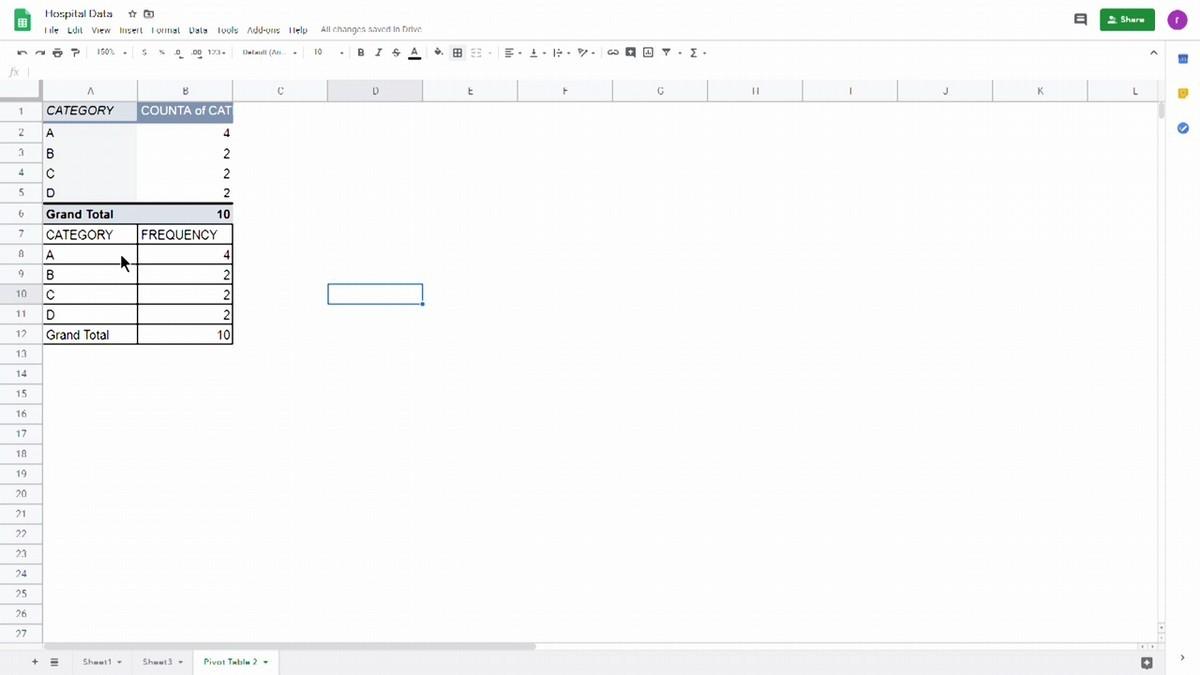
अब, आइए एक अंतिम उदाहरण देखें जहां मेरे पास इतने सारे अवलोकन हैं। फिर से अगर मैं एक मिलान चिह्न A, A, B, C, A, D, A, B, D, C, A, B, C, D, A, C, D, D करती हूं। तो अगर मैं इसे देखती हूं, मेरे पास 6 है, मेरे पास 3 है, मेरे पास 4 है, मेरे पास 5 है। तो, मेरे पास यहां 18 अवलोकन हैं। अब, यदि आप इसे देखें, तो मेरे पास यहाँ एक मिलान चिह्न है। मेरे पास 3 था, मेरे पास 6 था, मेरे पास एक था, फिर मेरे पास 1, 2, 3 था, मेरे पास 3 था। मेरे पास यही है। इस प्रकार, आप देख सकते हैं कि हम विभिन्न फ्रीक्वेंसी (frequency) बंटन बना सकते हैं।

(Refer Slide Time: 8:48)



गूगल शीट (Google sheet) में फ़्रीक्वेंसी टेबल (frequency table) कैसे करूँ? तो, आइए देखें कि गूगल शीट (Google sheet) में फ़्रीक्वेंसी टेबल (frequency table) कैसे करें। मैं वही उदाहरण देखती हूं जो मैंने लिया है।

(Refer Slide Time: 9:00)



तो यह मेरे पास गूगल शीट (Google sheet) है। मैं निर्माण करने जा रही हूं, मैं यहां एक शीट (sheet) जोड़ने जा रही हूं। उस शीट (sheet) में, मैं पहले श्रेणी का नाम टाइप (type) करने जा रही हूँ। तो, मेरे पास यहां एक श्रेणी है। मेरे पास जो भी डेटा (data) है, मैं उसे लिखने जा रही हूं, मैं डेटा (data) पर सूचीबद्ध करने जा रही हूं। मेरे पास A है, मेरे पास A, B, C, A, D, A, B, D, C है। तो, आप देख सकते हैं कि यह पहला उदाहरण है जिसे हम देखते हैं, मैंने यहां डेटा (data) सूचीबद्ध किया है।

तो, आप चरण एक पर वापस जाएं, आपके पास मौजूद सेल (cell) को हाइलाइट (highlight) करें, मेरे पास कौन से सेल (cell) हैं, मेरे पास ये सेल (cell) हैं, मैं इन सेल (cell) को हाइलाइट (highlight) कर रही हूं। यह पहला कदम है। अब आप फ़ॉर्मेटिंग बार (formatting bar) में दूसरे चरण को देखें। डेटा (data) विकल्प पर क्लिक (click) करें। इसलिए, मैं फ़ॉर्मेटिंग बार (formatting bar) में जाती हूं, मैं डेटा (data) विकल्प पर क्लिक (click) करती हूं। फिर, डेटा (data) विकल्प में, पिवट टेबल (pivot table) विकल्प पर जाएं। पहले मैं अपना डेटा हाइलाइट (data highlight) करती हूं, मैं डेटा (data) विकल्प पर जाती हूं, मेरे पास वह है जिसे मैं पिवट टेबल (pivot table) विकल्प कहती हूं।

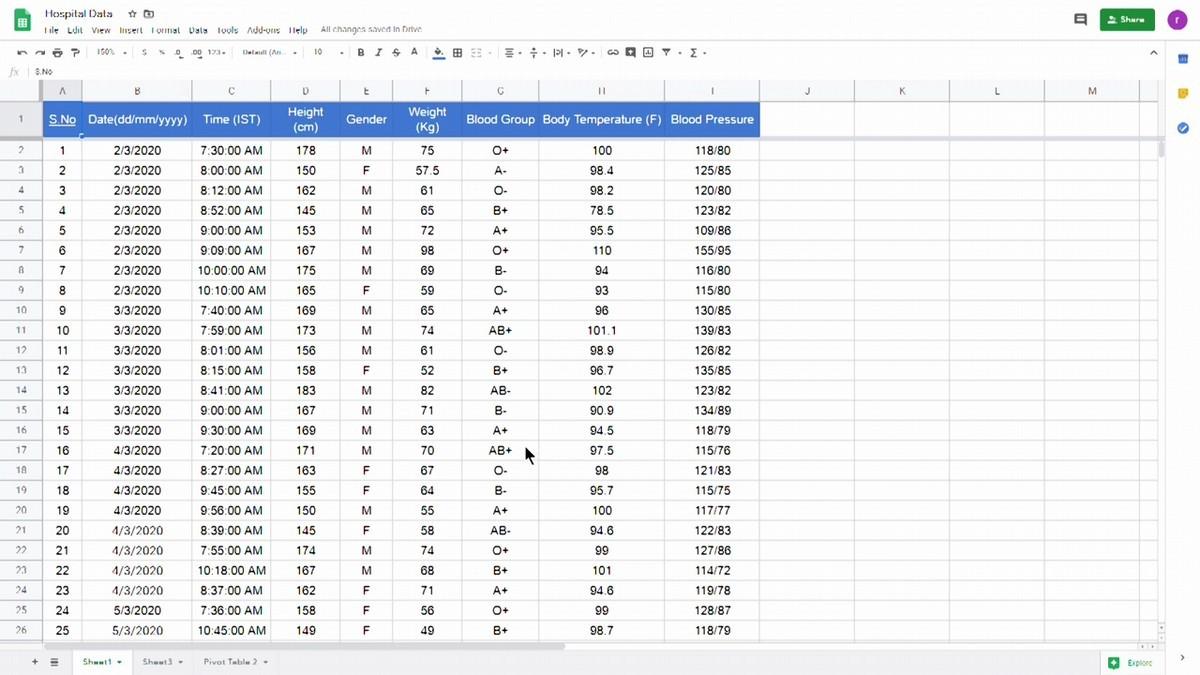
अब, यह पिवट टेबल (pivot table) में मैं सीमा निर्दिष्ट करती हूं। आप देखते हैं कि श्रेणी A एक से A ग्यारह तक निर्दिष्ट है जिसमें सेल (cell) A1 निर्दिष्ट करता है कि श्रेणी क्या है। यहां मैंने सिर्फ नाम श्रेणी दी है, मैं इस श्रेणी के वेरिएबल (variable) को कोई भी नाम दे सकती हूं। मैं एक अल्फाबेट (alphabet) दे सकती थी या मैं सिर्फ यह बता सकती थी कि यह कुछ समूह है, कुछ भी, लेकिन यह मैं सिर्फ एक निश्चित श्रेणी निर्दिष्ट कर रही हूं, इसे पिवट टेबल (pivot table) बनाने के लिए कह रही हूं।

अब, अंतिम चरण पर चलते हैं। पिवट टेबल एडिटर (pivot table editor) में पिवट टेबल (pivot table) बनाने के बाद, पिवट टेबल एडिटर (pivot table editor) क्या होता है, आपके पास पिवट टेबल एडिटर (pivot table editor) होता है जो दाईं ओर दिखाई देता है। वहां आप पंक्तियां जोड़ते हैं। पंक्ति में, मैं सिर्फ एक श्रेणी जोड़ती हूं। विभिन्न श्रेणियां क्या हैं? मेरे पास श्रेणी A, B, C, और D है। और मूल्यों में, मैं यह जोड़ने जा रही हूं जो कि श्रेणी में 4, 2, 2, 2 के मूल्य हैं।

तो, एक टेबल (table) बनाने का एक तरीका है, मैं बस इसे कॉपी (copy) कर सकती हूं और मैं मान पेस्ट (paste) कर सकती हूं। मैं यहां फ्रीक्वेंसी (frequency) के साथ एक श्रेणी तालिका दे सकती हूं और मैं देख सकती हूं कि यह कुछ और नहीं बल्कि तालिका है जिसे हमने अभी बनाया है। तो, यह आपकी गूगल शीट (Google sheet) में फ़्रीक्वेंसी टेबल (frequency table) बनाने का एक तरीका है।

इसलिए, एक बार जब आपके पास आपकी फ्रीक्वेंसी (frequency) तालिका, आपकी फ्रीक्वेंसी (frequency) तालिका होती है, तो हमने पहले उदाहरण 4, 2, 2, 10 के लिए ठीक यही किया है। आप देख सकते हैं कि हमारी गूगल शीट (google sheet) यही देती है, A की फ्रीक्वेंसी (frequency) 4, B की फ्रीक्वेंसी (frequency) 2 , C की फ़्रीक्वेंसी (frequency) 2, और D की फ़्रीक्वेंसी (frequency) 2 कुल 10 के साथ। यह वही है जो हमारे यहाँ है, पहली फ़्रीक्वेंसी टेबल (frequency table) जिसे आपने गूगल शीट (google sheet) पर बनाया है।

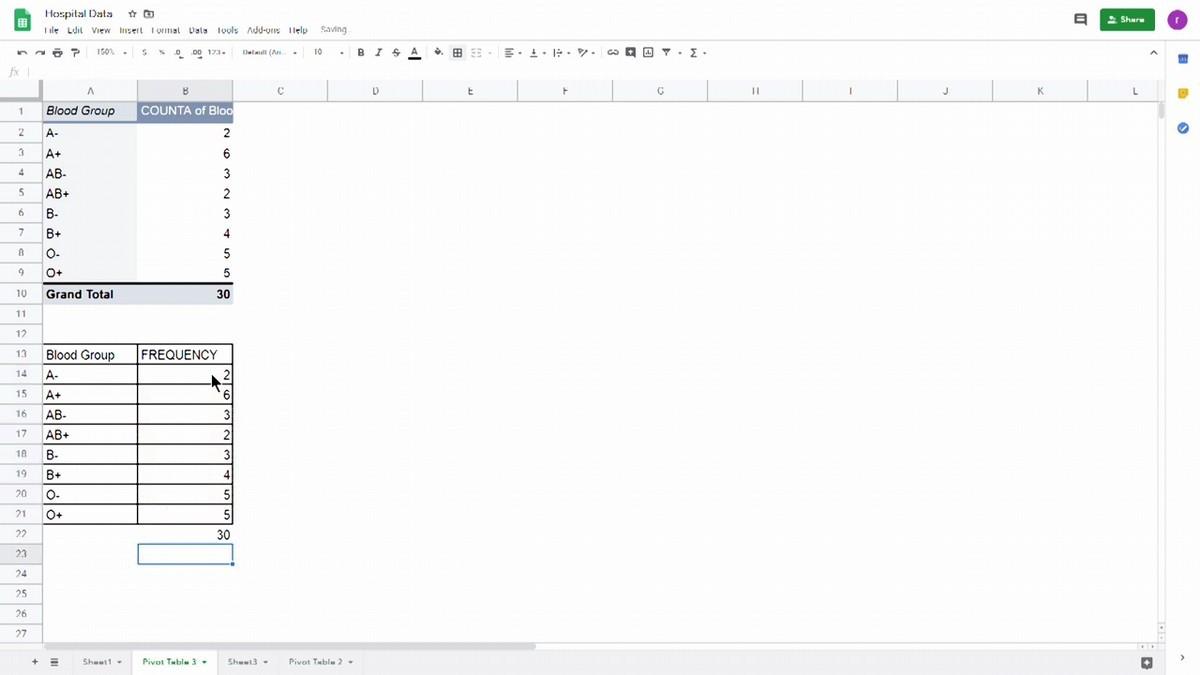
(Refer Slide Time: 12:49)



अब फिर से, मैं किसी दिए गए डेटा (data) के लिए वही तालिका बना सकती हूं। यदि आपको याद हो तो यह ब्लड ग्रुप डेटा (blood group data), अस्पताल डेटा (data) है जिसकी चर्चा हमने अपनी पिछली कक्षा में की थी। मान लीजिए कि मैं यहां एक विशेष श्रेणीबद्ध वेरिएबल (variable) के वितरण को जानना चाहती हूं, मैं यहां दो श्रेणीबद्ध वेरिएबल (variable) देख सकती हूं। एक जेंडर (gender) और दूसरा ब्लड ग्रुप (blood group)। जब मैं ब्लड ग्रुप (blood group) को देखती हूं, तो मेरे पास यही है। यह श्रेणीबद्ध वेरिएबल (variable) है।

इसलिए, मैं अपने पिवट टेबल (pivot table) चरण पर वापस जाती हूं। तो, मुझे क्या करना चाहिए, मुझे याद है, पहले मैं सेल (cell) का चयन करती हूं, मैं डेटा (data) विकल्प पर क्लिक (click) करती हूं, मैं पिवट टेबल (pivot table) विकल्प पर जाती हूं और मैं पिवट टेबल एडिटर (pivot table editor) के पास जाती हूं। तो, हम यहां वही काम करने जा रहे हैं। मैं इस डेटा (data) का चयन करती हूं, मैं डेटा (data) विकल्प पर क्लिक (click) करती हूं। यहां मैं पिवट टेबल (pivot table) विकल्प पर क्लिक (click) करती हूं।

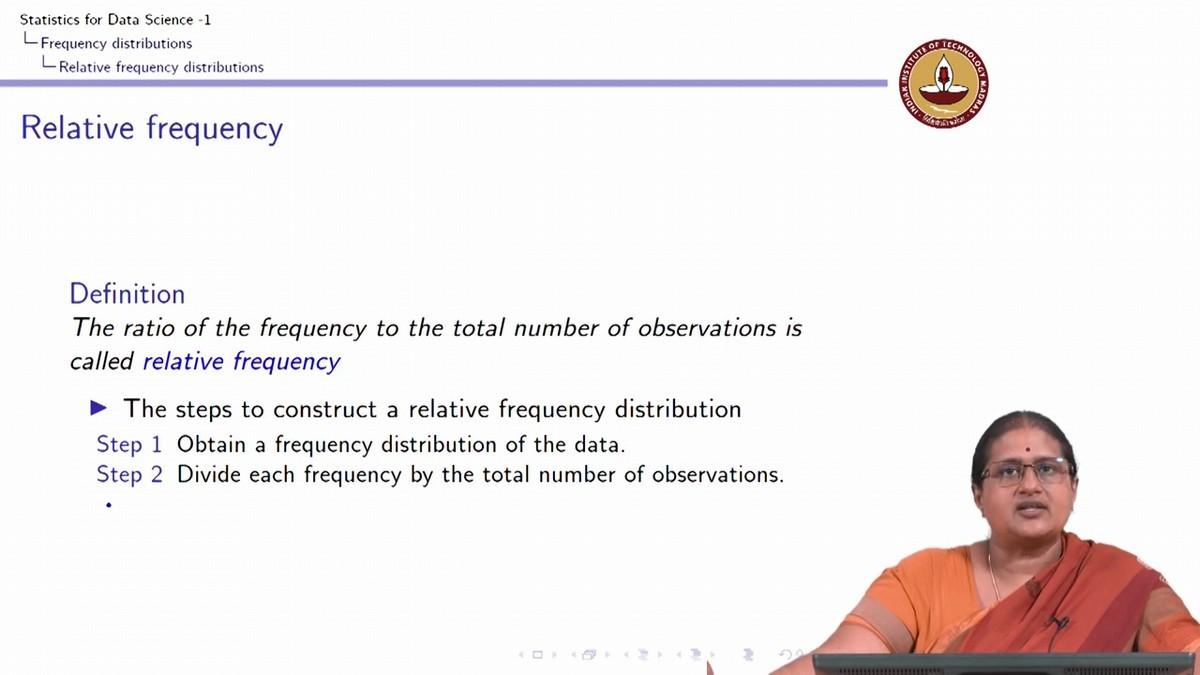
(Refer Slide Time: 13:47)



मैं एक नई शीट (sheet) बनाती हूं। एक बार जब मैं एक नई शीट (sheet) बना लेती हूं, तो मैं पिवट टेबल एडिटर (pivot table editor) पर जाती हूं। मैं पंक्तियों को जोड़ती हूं। अब आप वेरिएबल (variable) का नाम देख सकते हैं, अब यहाँ एक ब्लड ग्रुप (blood group) है। यही वह है जो मैं वेरिएबल (variable) ब्लड ग्रुप (blood group) के फ्रीक्वेंसी (frequency) वितरण को जानना चाहती हूं। इसलिए, मैं यहां ब्लड ग्रुप (blood group) पर क्लिक (click) करती हूं, और मैं मूल्यों पर जाती हूं और मैं मूल्यों को जोड़ती हूं और आप देख सकते हैं कि यह ब्लड ग्रुप (blood group) का फ्रीक्वेंसी (frequency) वितरण है। मैं सिर्फ कॉपी (copy) कर सकती हूं, मैं मूल्यों को पेस्ट (paste) कर सकती हूं। और मैं बस यहाँ कह सकती हूँ, यह ब्लड ग्रुप (blood group) है, और यह फ्रीक्वेंसी (frequency) है।

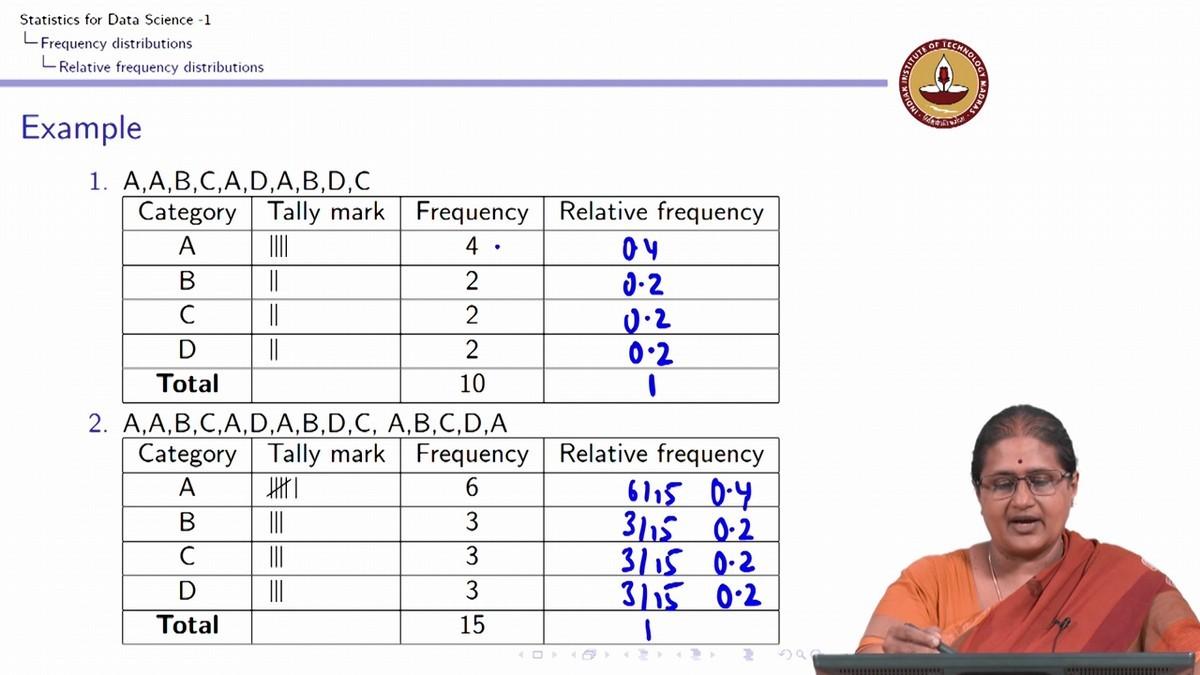
और आप देख सकते हैं कि यह मेरे ब्लड ग्रुप (blood group) की तालिका का फ्रीक्वेंसी (frequency) वितरण है, जो मुझे गूगल शीट (Google sheet) में मिलता है। यदि आप योग को देखें, तो आप उन सभी अध्ययनों का योग देख सकते हैं और ठीक यही मेरे पास 30 अवलोकन हैं, यह ब्लड ग्रुप (blood group) है जिसमें 30 लोगों की फ्रीक्वेंसी (frequency) होती है। तो, इस प्रकार हम मैन्युअल (manual) रूप से फ़्रीक्वेंसी टेबल (frequency table) का निर्माण करते हैं जो कि टैली मार्क्स (tally marks) का उपयोग करने के पहले सिद्धांतों के माध्यम से होता है और यह गूगल शीट (Google sheet) का उपयोग करके होता है। फ़्रीक्वेंसी (frequency) तालिका प्रत्येक वेरिएबल (variable), प्रत्येक श्रेणीगत वेरिएबल (variable) की गिनती देती है।

(Refer Slide Time: 15:43)



एक और चीज है जो बहुत उपयोगी है और वह है सापेक्ष फ्रीक्वेंसी (frequency)। आपेक्षिक फ़्रीक्वेंसी (frequency) जो कैप्चर (capture) करती है वह प्रेक्षणों की कुल संख्या के लिए फ्रीक्वेंसी (frequency) का अनुपात है।

(Refer Slide Time: 15:59)

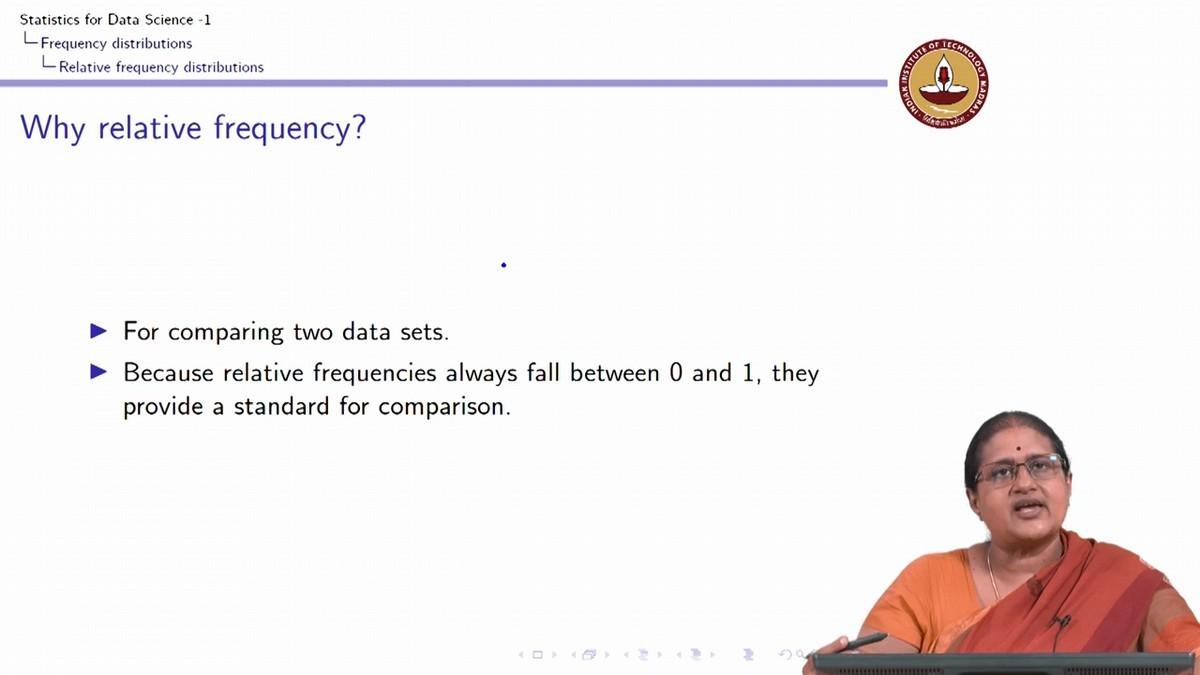


तो, हमने पहले से ही एक फ्रीक्वेंसी (frequency) तालिका का निर्माण किया है, हमने पहले ही देखा है कि 4 A, 2 B, 2 C और 2 D हैं। फ्रीक्वेंसी (frequency) का अनुपात, इसलिए, 4 A की फ्रीक्वेंसी (frequency) है, अवलोकन की कुल संख्या है 10 I जो अनुपात 4/10 है, जो 0.4 है, मुझे इस तालिका में A की सापेक्ष फ्रीक्वेंसी (frequency) देता है। तो, सापेक्ष फ्रीक्वेंसी (frequency) वितरण में, मैं केवल प्रत्येक फ्रीक्वेंसी (frequency) को अवलोकनों की कुल संख्या से विभाजित करती हूं, और मुझे 0.4, 0.2, 0.2, 0.2 मिलता है। और एक सापेक्ष फ्रीक्वेंसी (frequency) का योग हमेशा 1 तक जोड़ना चाहिए।

अब, यदि आप इस डेटा (data) को देखें, तो यह 0.4, 0.2, 0.2, और 0.2 था और 1 तक जोड़ रहा था। यहाँ आप इसे देखें, यह 6/15 होने वाला है, यह 3/15 होने वाला है, यह 3/15 होने जा रहा है, यह 3/15 होने जा रहा है। और आप देख सकते हैं कि उन सभी का योग 15/15 है, जो कि 1 है। आप फिर से देख सकते हैं कि यह एक 0.4 है, यह 0.2 है, यह 0.2 है, यह 0.2 है।

अब, जो मैं आपको देखना चाहती हूं, वह यह है कि इन दोनों वितरणों की फ्रीक्वेंसी (frequency) अलग-अलग है। यहां, मेरे पास A था जो 4, 2, 2, 2 है। यहां फ्रीक्वेंसी (frequency) 6, 3, 3, 3 थी। जबकि जब आप इस डेटासेट (dataset) और इस डेटासेट (dataset) की सापेक्ष फ्रीक्वेंसी (frequency) को देखते हैं, तो आप पा सकते हैं कि सापेक्ष फ्रीक्वेंसी (frequency) इस डेटासेट (dataset) में A की, श्रेणियों में से प्रत्येक की सापेक्ष फ्रीक्वेंसी (frequency) या A, B, C, और D में से प्रत्येक वेरिएबल (variable) इस डेटासेट (dataset) में A, B, C और D की सापेक्ष फ्रीक्वेंसी (frequency) के समान है।

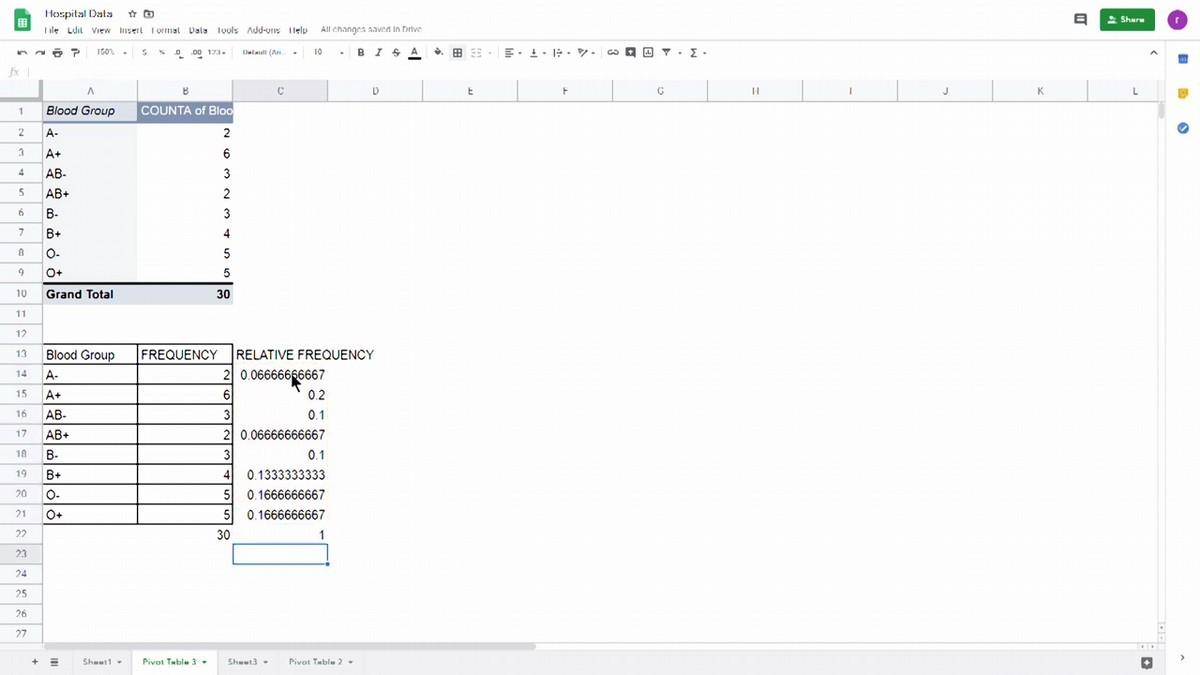
(Refer Slide Time: 18:07)



अब, हमें सापेक्ष फ्रीक्वेंसी (frequency) की आवश्यकता क्यों है, इसका कारण मैंने यहां प्रदर्शित किया है, भले ही गिनती में इन दो डेटासेट (dataset) के बीच अंतर है, आप देख सकते हैं कि सापेक्ष फ्रीक्वेंसी (frequency) काफी समान है। यहां, मेरे पास पूरी तरह से 10 अवलोकन थे, यहां मेरे पास पूरी तरह से 15 अवलोकन थे, लेकिन दोनों में एक ही सापेक्ष फ्रीक्वेंसी (frequency) है, फ्रीक्वेंसी (frequency) अलग है।

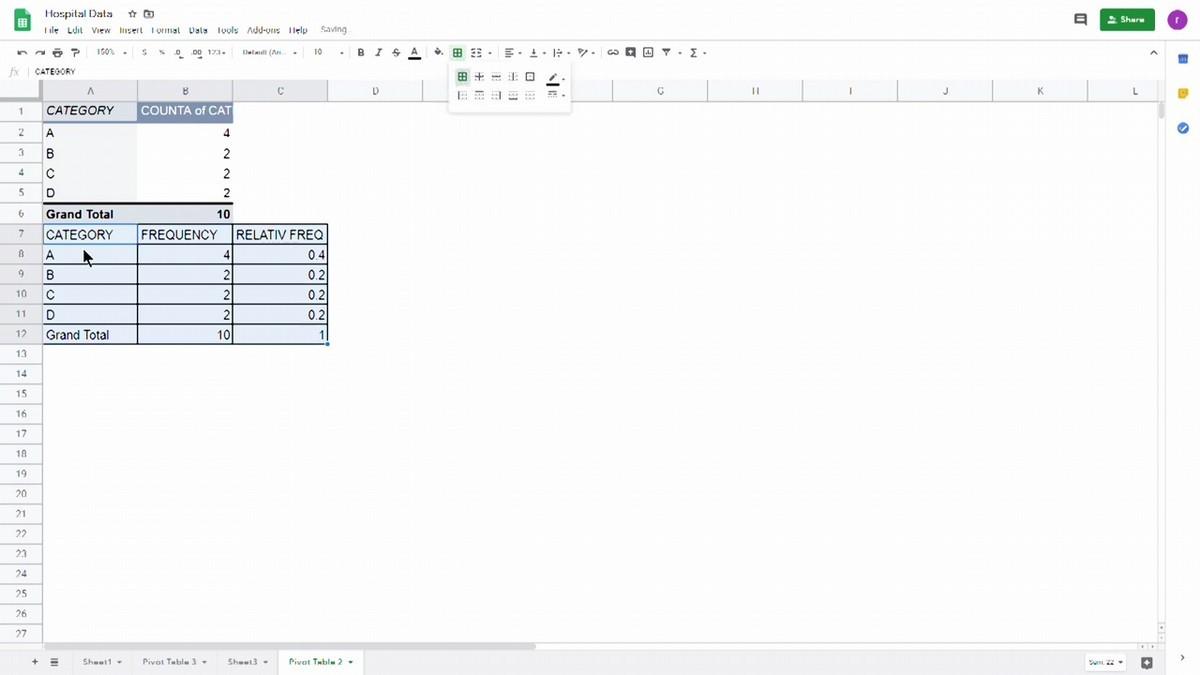
तो, सापेक्ष फ्रीक्वेंसी (frequency) हमें मदद करती है, दो डेटासेट (dataset) की तुलना करने में। और क्योंकि सापेक्ष फ्रीक्वेंसी (frequency) हमेशा 0 और 1 के बीच होती है, यह तुलना के लिए एक अच्छा मानक है। इसलिए, हम हमेशा एक सापेक्ष फ्रीक्वेंसी (frequency) तालिका रखना पसंद करते हैं।

(Refer Slide Time: 19:06)



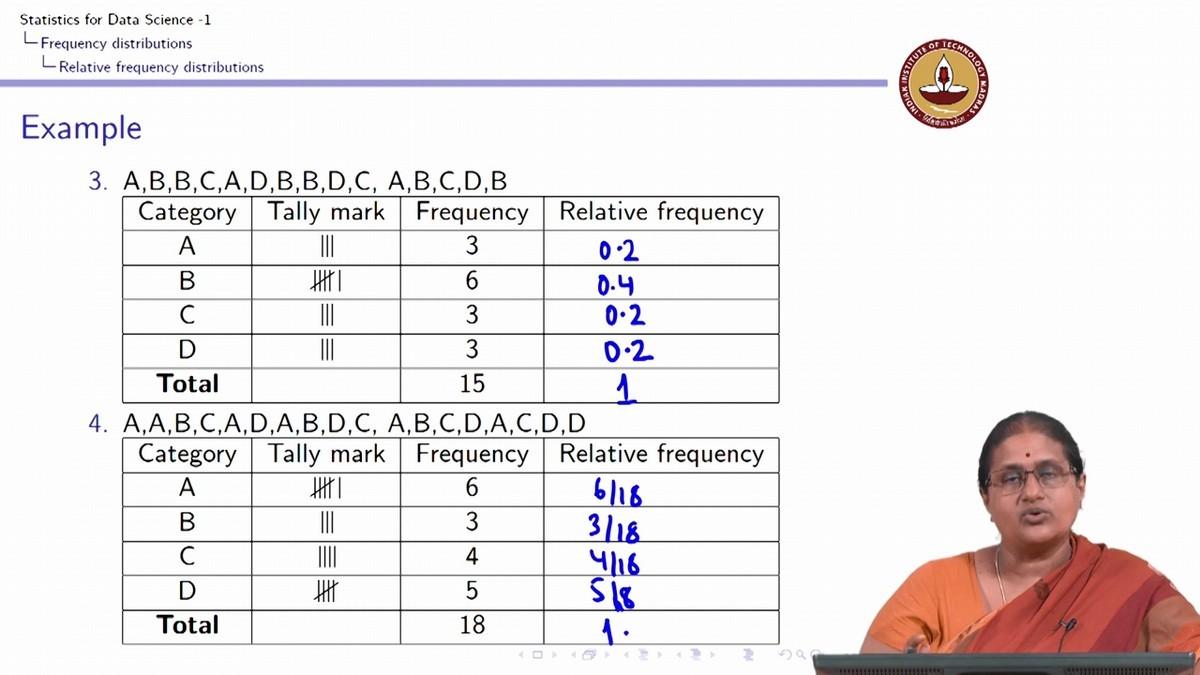
हम गूगल शीट (Google sheet) में एक सापेक्ष फ्रीक्वेंसी (frequency) तालिका कैसे बनाते हैं? गूगल शीट (Google sheet) में, हमारे पास पहले से ही एक फ़्रीक्वेंसी टेबल (frequency table) है। मैं यहां एक सापेक्ष फ्रीक्वेंसी (frequency) कॉलम (column) बनाती हूं, और मुझे पता है कि सापेक्ष फ्रीक्वेंसी (frequency) कुल से विभाजित फ्रीक्वेंसी (frequency) के अलावा और कुछ नहीं है। इस तरह मैं इसे परिभाषित करती हूं। और मैं इसे नीचे खींच सकती हूं। और आप देख सकते हैं कि, अगर मैं इन मूल्यों के योग को देखती हूं, तो यह 1 तक जुड़ जाता है, इनमें से प्रत्येक फ्रीक्वेंसी (frequency) मुझे सापेक्ष फ्रीक्वेंसी (frequency) देती है। यह ब्लड ग्रुप (blood group) के लिए है।

(Refer Slide Time: 20:03)

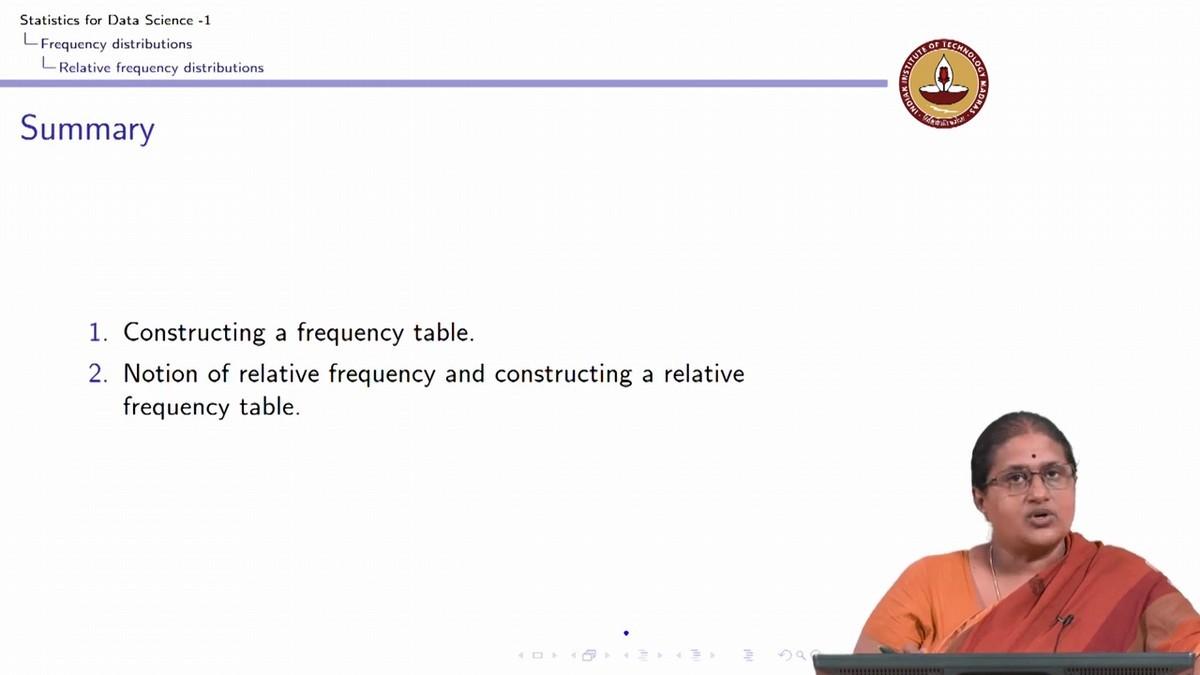


मैं इस पिवट टेबल (pivot table) के लिए वही काम कर सकती हूं। मैं सापेक्ष फ्रीक्वेंसी (frequency) कर सकती हूं, जो कि केवल 0.4, 0.2, 0.2, 0.2 के बराबर है, और सभी सापेक्ष फ्रीक्वेंसीयों (frequency) का योग हमेशा 1 तक जुड़ जाएगा। तो, अब मेरे पास यह है, जो मुझे देने वाला है, यह ब्लड ग्रुप (blood group) नहीं है, यह कैटेगरी (category) है। तो, मेरे पास पहले उदाहरण के लिए है, मेरे पास है जिसे मैं फ्रीक्वेंसी (frequency) और सापेक्ष फ्रीक्वेंसी (frequency) को श्रेणी वेरिएबल (variable) के साथ सूचीबद्ध करती हूं।

(Refer Slide Time: 20:56)



तो, मेरे पास दो और उदाहरण हैं। पहले के उदाहरणों में, तो, यह 3/15 होने जा रहा है, जो कि 1 बटा 5 है, इसलिए मेरे पास 0.2 , 6/15 है, जो 2/5 है, जो 0.4 है। मेरे पास फिर से 0.2 है, मेरे पास 0.2 है। तो, उनमें से प्रत्येक के लिए, आप देख सकते हैं कि यह 1 तक जुड़ जाता है। मैं इसे एक अभ्यास के रूप में छोड़ती हूं, लेकिन आप देख सकते हैं कि यह 6/18 होने वाला है, यह 3/18 होगा, यह है 4/18, और 5/18 होने जा रहा है। आप देख सकते हैं कि यह 5+4, 9, 9+ 9, 18 जोड़ता है। तो यह फिर से 1 हो जाता है।

(Refer Slide Time: 21:35)

तो संक्षेप में, हमने इस भाग में जो सीखा है वह यह है कि फ्रीक्वेंसी (frequency) तालिका कैसे बनाई जाती है? सापेक्ष फ्रीक्वेंसी (frequency) की अवधारणा क्या है? और आप एक सापेक्ष फ्रीक्वेंसी (frequency) तालिका कैसे बनाते हैं?

**Glossary:**

| **English Word** | **Word** | **Meaning** |
| --- | --- | --- |
| Lecture | लेक्चर | व्याख्यान |
| Statistics | स्टेटिस्टिक्स | अंक-विवरन |
| Sample | सैंपल | नमूना |
| Population | पापुलेशन | जनसंख्या |
| Structured Data | स्ट्रक्चर्ड डेटा | संरचित डेटा |
| Data | डेटा | डेटा |
| Variable | वेरिएबल | चर |
| Column | कॉलम | कॉलम |
| Cross-sectional Data | क्रॉस-सेक्शनल डेटा | क्रॉस-सेक्शनल डेटा |
| Nominal Variable | नॉमिनल वेरिएबल | नाममात्र चर |
| Ordinal Variable | ऑर्डिनल वेरिएबल | सामान्य चर |
| Module | मॉड्यूल | भाग |
| Association | एसोसिएशन | संघ |
| Frequency | फ्रीक्वेंसी | आवृत्ति |
| Tally Marks | टैली मार्क्स | अंकों का मिलान |
| Tallies | टैलीस | गणना |
| Flip | फ़्लिप | फेर बदल |
| Google Sheet | गूगल शीट | गूगल शीट |
| Frequency Table | फ़्रीक्वेंसी टेबल | आवृत्ति तालिका |
| Type | टाइप | टाइप |
| Cell | सेल | कक्ष |
| Highlight | हाइलाइट | उभारना |
| Formatting Bar | फ़ॉर्मेटिंग बार | स्वरूपण बार |
| Click | क्लिक | क्लिक |
| Pivot Table | पिवट टेबल | पिवट तालिका |
| Alphabet | अल्फाबेट | वर्णमाला |
| Pivot Table Editor | पिवट टेबल एडिटर | पिवट टेबल एडिटर |
| Paste | पेस्ट | पेस्ट |
| Blood Group | ब्लड ग्रुप | रक्त समूह |
| Copy | कॉपी | कॉपी |
| Capture | कैप्चर | दर्शाना |
| Dataset | डेटासेट | डेटासेट |
| Category | कैटेगरी | श्रेणी |