03. Data

ID 413: Information Graphics and Data Visualization Spring 2025

Venkatesh Rajamanickam (@venkatrajam) venkatra@iitb.ac.in http://info-design-lab.github.io/ One of the most important questions that a visualization attempts to answer is, "Compared to what?" Comparisons are at the heart of understanding quantitative information.

To help with easy comparison, we do one or more of the following:

- 1. Translate
- 2. Transpose
- 3. Transform

Translate -- Apply appropriate measurement scales to comprehend the data. Raw numbers to percentages, time into years, months, days, hours, minutes or seconds, lengths/distances into cms, mts or kms, areas into sq.units, volumes, weights, temperatures, and so on.

How many people are there in the world? There are almost 8 billion of us and we're scattered all over the planet! What is life like for everyone? Do you think it's the same? It's tricky to picture 8,000,000,000 people, so instead, let's imagine the whole planet is a village where 100 people live.* Welcome to our global village! Come and meet the people who live here. Each person represents 80 million people [more or less] in the real world. How many children and grown-ups are there? Are there more males or females? Let's find out . . . 50 26 people out of 100 are people are 50 children under 14 years old female people are male 8 people out of 100 are more than 65 years old If we focus on 100 people, it's easier to *The people we show in our village do not represent the actual age, gender or ethnic background of see the things we have in common and people in the real world, except on these two pages the things that make us different. where we do look at gender and age.

Transform -- Convert measurement scales into something more suitable for the visualization goal. Make it familiar & relatable to the reader.

Filtering/Refining: Removing irrelevant data points. Replacing codes with descriptive labels.

Aggregation: Combining data points into summaries like averages or totals

Normalization: Scaling data to a common range. Negative & log scales.

Binning: Grouping data into discrete ranges. For example, grouping age ranges into "young adult," "middle-aged," and "senior".

Data type conversion: Changing data types (e.g., text to numeric)

Transform -- Convert measurement scales into something more familiar & relatable to the reader.



TERMS OF SERVICE

HOW LONG IT WOULD TAKE TO READ THE TERMS OF SERVICE AGREEMENTS OF POPULAR ONLINE SERVICES

Even the shortest terms and conditions for popular online services are a few thousand words long.

10.0

As a result, 97% of people in the 18-34 age group agree to conditions without reading them.

10 10 10

Reading time Assuming 240 WPM

Length of text Total word count

G	0	8	0	in	0	D	Ś	a	0	5	N		Uber
PACEBOO		SPOTER	TWITTLR	LINKEDIN	TINDER	YOUTUBE	APPLE	AMAZON	ZOOM	TIKTOK	NETFUX	MCROSOFT	LEER
						-				-			
			-	=	=		=		=	=			
				E	5								
							-						
			<u>_</u>			-				-			
		-					1						
							Ē						
	and the second se				the second se		the second se					-	_

Transpose -- For impact & surprise, transposition data in unexpected ways.



Translate -- Apply appropriate measurement scales to comprehend the data. Raw numbers to percentages, time into years, months, days, hours, minutes or seconds, lengths/distances into cms, mts or kms, areas into sq.units, volumes, weights, temperatures, and so on. Negative & log scales.

Transform -- Convert measurement scales into something more familiar & relatable to the reader.

Transpose -- For impact & surprise, transposition data in unexpected ways.

In this assignment, take a very small dataset and create 3 visualizations by applying translation, transformation and transposition of the data. Post your submission as a comment in this issue.

Approximately how many accidental deaths occur in the Mumbai suburban train network (Mumbai Local)?

- a. 1200
- b. 1800
- c. 2500
- d. 3500

What is Mumbai's approximate current population?

- a. 18 million
- b. 22 million
- c. 24 million
- d. 26 million

During peak hours, how many people pack into one compartment of a typical Mumbai local train?

- a. 300
- b. 400
- c. 500
- d. 600

During peak hours, how many people pack into one compartment of a typical Mumbai local train?

- a. 300
- b. 400
- c. 500
- d. 600

Do you know how much it is in terms of number of people per square meter?

Do you know the actual capacity of a compartment?

During peak hours, how many people pack into one compartment of a typical Mumbai local train?

- a. 300
- b. 400
- c. 500
- d. 600

Do you know how much it is in terms of number of people per square meter? (14-16 per sq.mt)

Do you know the actual capacity of a compartment? (188)

How many cases are pending in various judicial courts of Mumbai?

- a. 3 lakhs
- b. 6 lakhs
- c. 12 lakhs
- d. 24 lakhs

How many cases are pending in various judicial courts of Mumbai?

- a. 3 lakhs
- b. 6 lakhs
- c. 12 lakhs
- d. 24 lakhs

Do you know how of them are criminal cases?

How many cases are pending in various judicial courts of Mumbai?

- a. 3 lakhs
- b. 6 lakhs
- c. 12 lakhs
- d. 24 lakhs

Do you know how of them are criminal cases? (40,000)

What is the car ownership in India?

- a. 18 per 1000
- b. 34 per 1000
- c. 58 per 1000
- d. 72 per 1000

What is the car ownership in India?

- a. 18 per 1000
- b. 34 per 1000
- c. 58 per 1000
- d. 72 per 1000

Do you know that this is lower than Sudan & Afghanistan

China? Europe? U.S?

What is the car ownership in India?

- a. 18 per 1000
- b. 34 per 1000
- c. 58 per 1000
- d. 72 per 1000

Do you know that this is lower than Sudan & Afghanistan

China? Europe? U.S? (58, >500, >800)

What is number of tax payers in relation to voters?

- a. 4%
- b. 7%
- c. 12%
- d. 18%

How much did India's Mars programme cost?

- a. \$44 million
- b. \$54 million
- c. \$64 million
- d. \$74 million

How much did India's Mars programme cost?

- a. \$44 million
- b. \$54 million
- c. \$64 million
- d. \$74 million

NASA's Maven mission – \$ 671 million ESA Mars mission – \$ 386 million Japan's failed Mars mission – \$ 186 million

How much did India's Mars programme cost?

- a. \$44 million
- b. \$54 million
- c. \$64 million
- d. \$74 million

NASA's Maven mission – \$ 671 million ESA Mars mission – \$ 386 million Japan's failed Mars mission – \$ 186 million

Rs. 7 per km

How much did India's Mars programme cost?

- a. \$44 million
- b. \$54 million
- c. \$64 million
- d. \$74 million

NASA's Maven mission – \$ 671 million ESA Mars mission – \$ 386 million Japan's failed Mars mission – \$ 186 million

Rs. 7 per km (Mumbai auto Rs. 14.20 per km)

Ratio of fake doctors to real doctors practicing in Delhi:

- a. 1:4
- b. 1:3
- c. 1:2
- d. 1:1

Intelligence Hierarchy: Data, Information, Knowledge, Wisdom



Nathan Shedroff



David McCandless // v 0.1// work in progress InformationIsBeautiful.net Measurement is the foundation of all data

"Whenever you can, count." — Sir Francis Bacon

"One accurate measurement is worth a thousand expert opinions." — Grace Hopper

"In many spheres of human endeavour, from science to business to education to economic policy, good decisions depend on good measurement." — Ben Bernanke

"I have been struck again and again by how important measurement is to improving the human condition."

— Bill Gates

"Measurement is the first step that leads to control and eventually to improvement. If you can't measure something, you can't understand it. If you can't understand it, you can't control it. If you can't control it, you can't improve it."

— H. James Harrington

Pre-modern customs were all about dealing with trust, the need for direct supervision, and facing up to the enormous risks posed by nature.

Once fundamental measurement problems were solved — involving time, distance, weights, and power, among others — it became possible to cheaply measure worker performance, input and output quality, and the role of nature, in areas of life that were unheard of before.

This ability to cheaply measure ushered in the world of modern institutions.

— The Institutional Revolution: Measurement and the Economic Emergence of the Modern World (2011) by Douglas W. Allen





- So what is it about soccer that makes it so hard to quantify?
- Or what makes cricket so easy to measure?

- So what is it about soccer that makes it so hard to quantify?
- Or what makes cricket so easy to measure?
- One obvious answer is the quantity of the units that can be easily separated and analysed.

- So what is it about soccer that makes it so hard to quantify?
- Or what makes cricket so easy to measure?
- One obvious answer is the quantity of the units that can be easily separated and analysed.
- In cricket it is innings, overs, balls, runs, wickets, catches and so on; in basketball it is points, assists, rebounds, steals, turnovers, etc., in tennis it is serves, volleys, aces, errors (forced, unforced) etc.
- For soccer, the only apparent unit to separate out is the 45 minute halftime mark. Changes in possession could be another measure, or number of passes, or shots at goal etc., but none are very useful.

(Lionel Messi: http://fivethirtyeight.com/features/lionel-messi-is-impossible/ Stephen Currry: http://peterbeshai.com/buckets/app/#/playerView/201939_2015)
Data Semantics are the real-world meanings of the data, such as whether the data represents temperature or height measurements. Focuses on organizing data that reflects the basic meaning, rather than solely or primarily on the relationships and attributes of the data.

Data Behaviour encompasses the trends, patterns, and shape of the data values relative to each other.

Data Type is the level or scale of measurement that describes the nature of information within the numbers assigned to variables.

Derived Data are created from the original data values. These include measures such as count, summation, or average, transformations such as binning, or the inclusion of additional, external data.

Reflections on how designers design with data. Alex Bigelow, Steven Druckery, Danyel Fishery, Miriah Meyer. AVI '14 Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, pp. 17-24.

Characteristics of data

Data Semantics are the real-world meanings of the data, such as whether the data represents temperature or height measurements. Focuses on organizing data that reflects the basic meaning, rather than solely or primarily on the relationships and attributes of the data.

Data Behaviour encompasses the trends, patterns, and shape of the data values relative to each other.

Data Type is the level or scale of rement that describes the nature of information within the ners assigned to variables.

The first two characteristics, semantics and behaviour, are inherent to data and are not decided or modified in the process of creating a visualization. They are, however, critical for a designer to understand.

The latter two characteristics are an important part of the design decision process when working with data. Specifically, designers have the freedom to create, define, and modify data structure and derived data.

Data Behaviour encompasses the
shape of the data values relativehds, patterns, and
Zn other.

Data Type is the level or scale of measurement that describes the nature of information within the numbers assigned to variables.

Derived Data are created from the original data values. These include measures such as count, summation, or average, transformations such as binning, or the inclusion of additional, external data.

Data Semantics are the real-world meanings of the data, such as whether the data represents temperature or height measurements.

- Colour coding (semantically resonant colours)
- Data views (Vietnam War memorial)

Data Semantics



Lin, Sharon, et al. "Selecting Semantically-Resonant Colors for Data Visualization." Eurographics, vol. 32, 2013, pp. 401–410.

Data Semantics



Birthdays in the U.S and in India, Gramener

Data Semantics



https://www.fastcompany.com/3062182/how-to-use-color-to-prove-your-point-from-a-data-viz-expert

an terater usan natural som a lave opplation and and a som a lave opplation and a some of the source STRUCTURE FOR THE STRUCTURE AND THE STRUCTURE AN Constant - Kith A (2017- KiteSt Wind David, Safety Constant) ALVERIDGEN UNDER STREET TEXAL DE LE COMPLEXITE DE LE COMP LEVELS STVL - VEHICLE FOR ALL COMPAREMENTS AND ADDRESS REVERALLICE (1936 CALLER GREET ON CONTROL ON ETTE PRETER VERSETERED VERTEED - DOWLDF SHELVE (VERSENW-WEISHOW) E. S. C. MORT, A STORT ADDRESS OF A STORY a fik bekaletin - Educati e nedale incanti nedari pagi nena-adhe nang-AD E CALING - FASHERING - COMEL CARON- MONITORINH - CANONINA STAN COERLY-LE ADMINI-KENKENEL-IENCHINESKIN FEILINEOMI-CP LANCETURG - WELTON WITTER - CHREETENCE - CHERTON - CHRENT REPORT AND MORE FOR MARKED OF MERING AND THE MERING AND TH VALD H PAINER ESTELPENERSKI MARKELE MOUNTS- BACERERE MOUNTAGE DEBUG TEERS - CANADELS HAVE - RELEASED - RELEASED - RELEASED - CANADELS - CANADELS - RELEASED - CANADELS ERTESTROOP, ESTERAN WINDERA- JANES D'ENANG-CALMITIKANGA-MAKADISTURKA-ELL STATES - MEDING REPORT ON EACO, DURING AND AND AND A STATES NAS E HILL - ROBERT LOE VOE - CHALESDOM - KOHKO W DWOAN ENREKEERED ERT C. L. FEROLOGIUN + CLINIONE HARIOS - DAVID A HARD- HARAON HEIDERON- HOMAS (ULB) CHV WHINES HENRY BUCKNELLY DEVOLUND YOM ALD CHUNCHER-I M GAPELIOL DO WID WITE HARP C BORL HAR DURING COME DA ENT MATTERNA I A DENKO A VENA A SUCA DATARA (MATTANEN) A METROPAKA OVH STUDY - TEXTLE CARE - ROPE CARE - BOW OWNER WEIGHT TROON DILLE - EMIN WORK WLIE COMONIN- ENERGIAMED ONE ND M HOUD - 1/153 HOURD - NUCH HUSE - WIDE - HEDGE PER CONSA + CONSA : ANTICA MARKED

18 ALC: NOT THE OWNER. **CALLED AND AND** NAMES OF A DESCRIPTION OF A DESCRIPTIONO BALLING AND ADD COMMAND DESCRIPTION OF THE PROPERTY AND ADDRESS KARTIN STELLORI I REGISTORE HERE REGISTO ANTER THE REPORT OF A DESCRIPTION CHRONE-CALCER AND SWARTHARD (And as not as converse conversion and a converse of the conve SECOND ADD SOT CREMENDED ENGINEERS SERVER ALORA OF ALL MERIE-OFF MERIE-IN REPORTED AND TAKEN AND A DALARD THE PLACE OF A COMPANY OF THE PLACE DATINGS LEAR OF GROUP, NO ASSOCIATION DOGNATION OF A DESCRIPTION OF THE EXCENSION CONTRACTOR OF And Data and These lives and measures ends this made for performing ordinacionale Coloris da a la companya da a sua perso Mark Construction Distances and a new STATE OF CERCERO.

OCT IN CASE OF L

BY KING + DONALD H MOSES + ELVIN PRICE IT IN TETREAULT • HOMER E TISCHLER • KENN ERRILL V BROWN + MICHAEL T FITZPATRICK + DANNY G KEATON + ROTEPT MARKING KENNETH C KRAUS+JERRY LEE LEONARDI • FRED A LIND • HANS P E JOLLY RT P. SEJUSA · JOHNNY RAY TAGMAN · JAMES Q AYDLETT · EU SFERRAZZA ANN · RAMIRO MANRIQUE Jr · DAVID E NANCE · AN STEPHYSE WITHIN A LEON + BRAD W MORAND + WILLIAM L ROBERTSOF CHARLES & WILLIA GEORGE 5 YOHNINSON + MICHAEL G BARTON + RICHARD FEINKOWSKI + FELIPE VILLA WATHEN + DAVID REDATE + MARSHALL F KIPINA + ROBERT G NOPP + RICHARD G SCHROL ET + DONALD E WOOTBRICHT PAUL CHENDRIX • ORSON H CASE • NARCISO REZA CERSION MAIAI - PAUL R. CHAMBER MES W CHERRICK + JAMES J CONNELL + ROBERT J CROCE TON - RICHARD W DEMERS + MARK V DENNIS + WILLIAM L DOUC AS In LERS · MICHAEL A CUNI DDEN + ANTHONY G CALEFE LEY + RICHARD S HOWIE + JAMES L LEDIN + WILLIAM J LILLY . THOMAS A MCCONALTY SAMUELL REED . JAMES M REID . RALPH V RENAULD JL . TOESIER • HEROLIN T SIMMONS • LYNN SPANN • STIPHEN S STRYCHARZ IF • GERALD E STUT TAD ELFER • ISAIAH H WILSON • JOLEN D CAREY • CHARLES D CROWE • NORMAN E DAWSON J ODSEY · ROY D HEASTER · WIELFAM F LABRECQUE IF · RICHARD A MAUL · RICHARD S ME MULLEN / LARRY E POTTER · JESSE B STRIBLING · LLOYD B HASKELL Jr · JAMES L HOGGA IORT · JAMES D NEWCOMB · CHIRLES E SMITH · STANLEY R SMITH · ANTHONY TREVISE ANASIEWICZ + MICHAEL J BANOVEZ Jr + LEWIS C BARNARD + GARY W BARTLOW + RAFAEL & BEECH 3UTLER IF • ULYSSES C COOPER • RAMON DESCHAMPS • JOSEPH W DUNN • PATRICK P ... DWARDS - RONALD LEUNICE + RONALD E HOEL + JOHN T KYSER + RUDOLPECE RRS + PETE M MARTINEZ + DAVID E PORTERFIELD + DONALD J PRIEST + MICHAELP SMITH OMAS • THOMAS M WILLIAMS • CHARLES R MALBROUGH SHEARNE W BEAVER STAMES RESCHE TTESCH • KENNETH M BRANAMAN • DANNIE J BREWINGTON • JOHN T DOIKE • TH CLAWSON + CHARLES H CRANE + JAMES R CROLEY + JESUS DE LAROSA Jr + TERRY A / DIAMOND + BERNARD DILLARD + JACK M BROWN JF + CHARLES R FLEMING + CHARLES L GELMAN OSADO-RODRIGUEZ • ROBERT W RUHL • DALLAS E GREEN • BOBBY ROY HOLLEY • DAVID W C HUBBARD • LAMOND J JACKSON • ALLEN L JELINEK • RAYMOND E JOHNSON • THOM h Knott • Johnny F Long • Leroy Barnes • Michael J Mummel • Charles a mcguin NAGATO • JOSEPH E NEWMAN • LEROY E PETERSON • PATRICK T QUINN • JAMES M RADZI THE + SIM H RIGGINS Jr + JOSE C GOSSE + BERNARD A GREEN + TURNER L THOMPSON Jr + H WILSON IF · DARRYL G WINTERS · STEPHEN A ZUKOV · JAMES F ASKIN · CLIFFORD S BRATCHER (RNE IF • STEPHEN M CAMPBELL • S W GEORGE • TIMOTHY S DAVIES • DAVID A DILLON • ASLEY · RONALD M FILKINS · CLIFFORD L CARPENTER · BRENT I GRIGGS · EDWARD F HAP · IOWELL • LESTER JONES • ROLF W JORGENSEN • BILLY D. NELSON • DOUGLAS M KYSER • VMBDIN • MERRILL R LEWIS JF • ERNEST D MITCHELL • RONALD J KINKEADE • WILLIAM H NELSON • CHOLS • CRAIG R NOBERT • BRADLEY A PEARSON • DAVID E PETERS • CHARLES S RIDOUT • JTHERFORD • CHARLES W SMITH • KENT A SMITH • WALTER & TATE Jr • DONALD C WOODRUFF • RRY · RICHARD BRIGHT · WILLIE & CARY · STEPHEN K DIBB · DOUGLAS R FRENCH · CH + RICHARD D GILL Jr • THEODORE GRONOWSKI Jr • DENNIS L HARMON • SAMUEL G HARRIS • RLONG • PEDRO MARTINEZ • JOHN D TYLER • ROBERT E PULLIAM Jr • WILBERTO CABRERA SANCHEZ • DLTENBURG + RAINFORD TIFFIN + RICHARD F MENICHOLS + LARRY VAN CLIEF + OSCAR C BELL Jr + RRY · JAMES W BETHEL · ROY M BROOKS · WALTER 5 BROWN · EDWARD W BUTLER · MacLEAN + HENRY H DELANO + JOHN L DODSON + ROY FLORES MENO + MAXWELL S FRANTZ + MBLE • RICHARD A GILBREATH • JOHN C HOLOKA • RODNEY A JAMES • PAUL W JOHNSON • ANORE • JOHN J FLEMING • MITCHELL E COX • DORRIS E PATTON • JOHN W PERICH • esewonit • Phillip m Pierce Jr • John W Reagan • Juan Diego Sanchez • Chester J Simmons • MITH + ROBERT E STALLINGS + ROBERT J WELLS + RODNEY W WESTCOTT + GARLAND F WRIGHT + YOUMANS · FRED O BAUGH Jr · JAMES R BROWN · DONALD G DAVENPORT · MELVIN FORD · eary if • Daniel m gunn • David w Hill • Richard d Kaler • Jarvis C Lowder • MADISON + SAMUELT MORROW + BENJAMIN & NEWSOM + ROBERT LWENCER + STEVEN O

STELIMME DENNEY DONALD M SINGE REIDERE KEMP ALTEN G WELLS - JAMES E BU DERGEN A MAYER + LYMAN A MCMULLEN + EDW MEEVIN BARBER + ERIC & BARROW JE + DAVID E BRACKER DALE & CHARK + PETER DUGNESS I WITHOK IT FIKE - RONALD E HUMPHRIES - BILLY W MACHEN -CALL FAWKENCE D WATKINS JL FRANK J AUGUST MICHAEL J COMI LA MARE FISHER TRANK K LACHER + JOHN L HYMES + PAUL J KAPPMEYER + THOMAS T KASAL + MOMAS F KYTE IF WILLIE TOCKER / SILBANO MORIN / DANIELE MORRIS / SILEN C ROCK / DETZ · CARE W TUCKER · RICHARD == MILIKIN III · JOSEPH H WAETON · LARRY E WHITIN THOMAS A BLANEY . WILLIAM & DAVE · JAPALS R JOHNSON · MAITHEW J KRIST · REEBOUEF - FEEIX & RAMOS IF + WILFORD TO PRINCE S - HERMAN & ARMSTRONG + PETER W BA SOBBY G CORSE- HENRY V DEMPS - CHARLES E HEATH - METAELLEWIS - BILLY RAY OWENS -NORMAN C J POTRAS + JAMES L'STONE + ODIS THOMPSON + WALKER N WEEKS + LEROY F BARNES + PETER R BINGMAN + RONALD I BOWMAN + RONNIE LEE BRAY + JAMES R BURDETTE + FREDERICK K DEAN ANDRE L'DUBE - R'V'EDWARDS - VOLUE FRYER - KENNETH E GLAZE - PHILLIP H'CRECCE, DOUGLAS T'P WILLIAM L HARTLEY . DONALD LT DN + PAUL L HEATER + NORMAN A HEFT + CHARLES V PICCOLEL RAYFON LOFTON - WILLIAM B NOR 6 2011 - WILLIAM B HEPBURN + PETER QUILICED (+ PAUL M REED + KENNETH P TIERNEY ACTIFION M WALTER - DENNIS W WELLMANN - JOHN D. SANDERSON - CHARLES WAYNE R BAKER + DOUGLAS S DUBOSE - BILLY JOE HARRISON + JERRY W NYE + ROBERT L PERRY -ANTHONY LABRUZESE + STEPHEN F SNYDER + WIELIAM J TARSI + PETER H ALBRIGHT + TERRY LEE ALEXAN ALFRED J AMBURGEY + ADRIAN J ANGEIM + HERBERT S BECHTEL + JOEL BERNSTEIN + GYORGY J BESZE + RICHARD W BILLINGSEEY - RAYMOND A BROOKS - MAX L BRUBAKER - WILLIAM H BULLARD - DAVID E JOSEPH E CANDEAS - RICHARD D CLARK - THOMAS E CLARK - LEO C DIXON - JOHN F DOYLE -JAMES E DOZIER • OLIVER D DUMAS • LAURENCE J EXIMIN • JAMES W EMERSON • RUSSELL À ENCREN • STANLEY C CAMPBELL+ CARL L GLASSCOCK + DUANE I GREENLEE + CARL R HARBIN JE + VERNON G HOI FRANK P HUDAK + EEIAS M KAUHANE + LARRY R KEEM + ARTHUR G KLIPPEN + PETER S KNIGHT + SAMUEL J LARKIN / DANIEL C LYKINS · VERNON A MANHEIM Jr · RALPH E MASSEY · HAROLD E MATSON MATTHEW G MIEDEMA & ROBERT F MONTEITH IL: EDDIE O McBROOM Jr + RONALD P O'ROURKE +] C FELIX PACHECO + MAXIE PATTERSON + JOHN A RICH + BENJAMIN C ROLLER Jr.+ JULIUS M SANDERS + ALTER W SHIPLEY JE • FREDERICK B SKAGGS • FORTUNE SMITH • ROBERT L SMITH • WILLIAM M SMITH aymond m thompson - christopher c oakes - ronald L watson - james t beasley - john j DANIEL LEARLENBAUGH - JERRY WEDDY - RICHARD A FRYE - DONALD F. GIBSON - JULLS T. GIRTANNEL PEDRO R GURRERO - CHARLES E HUTCHINS JP> CEIFTION F VAUGHN - CARY M KISH - MARVIN D KOST onald lee lehew • Donald A lewis • James E reid • Arthur B Scarbrough • Donald V stei FIN W JOYS • CHARLES F WHITFIELD IF • ARTHUR P WILLIAMS • ELDRIDGE C AIGELDINGER • FRANCIS ROBERT L BANKS · RANDALL D HOWELL · EUGENE R MOREAU · KENNETH PARTLOW · WILLIAM D PERI RAYMOND J PONCURAK · EVERETT P RUNNELLS · STEPHEN J STEMAC · ALEXANDER THARP · LEO K TH MILTON L'VAUGHAN + JOHN E WATSON + ROBERT L'BABULA + WILLIAM R'TAYLOR + JOHN E BODENSCH ROBERT C BORTON IF * SAMUEL W BREWER * DENNIS R CARTER * THOMAS R COWLEY * WILLIAM A FI WILLIAM & KELLEY + PAUL & MADDEN + JOSEPH & MAMMOLITTI- JOSE LUIS MORENO + ARBAL ROLLIN DONALD B SMITH + FRANK L BARBEE + VICTOR M TOSADO-HERNANDEZ + JOHN R BALLARD + MARTIN VERNON E BRACKINS · CHARLES T MCCORKLE · RIGHARD A OSTEEN Jr · BRUCE M SMITH · WALTER L S WILLIAM & BUFF + PATRICK J CONNORS + JOHN T DEMALINE + SAMMIE DON HOFF + KENNETH D ROB LARRY D TAYLOR • ARTHUR E COMSTOCK Jr • JAMES B HEFFRON • ROBERT H LERNER • JAMES E METOY JAMES & TOOTHAKER + ALTON WILKINS III + MICHAEL CARROLL + LOWELL G EINARSON + GEORGE R KE JAY A HURD + HUBERT C NICHOLS IF + BRADFORD S NICKERSON + COLINNA FEEMSTER + NORMAN SC DAVID W ROBUSTELLINI · WILSON C BICKLEY · TERRENCE J CATHER · JOSEPH Q CONWAY · DANIEL J J HAROLD J KOLLER + EVEANS J STULTS + STANLEY G TOLER + ROBERT P ANDREWS + ROY J BULTMAN + THOMAS EMAILEDIN + CHARLES T DAY + DAVID C DEWITT + THOMAS E DORAN + DANIELS FERRY +

Data Type is the level or scale of measurement that describes the nature of information within the numbers assigned to variables.

It refers to the interpretation of an item in terms of scales of its measurement (attributes)

SCIENCE

Vol. 103, No. 2684

Friday, June 7, 1946

On the Theory of Scales of Measurement

S. S. Stevens

Director, Psycho-Acoustic Laboratory, Harvard University

FOR SEVEN YEARS A COMMITTEE of the British Association for the Advancement of Science debated the problem of measurement. Appointed in 1932 to represent Section A (Mathematical and Physical Sciences) and Section J (Psychology), the committee was instructed to consider and report upon the possibility of "quantitative estimates of sensory events"—meaning simply: Is it possible to measure human sensation? Deliberation led only to disagreement, mainly about what is meant by the term measurement. An interim report in 1938 found one member complaining that his colleagues by the formal (mathematical) properties of the scales. Furthermore—and this is of great concern to several of the sciences—the statistical manipulations that can legitimately be applied to empirical data depend upon the type of scale against which the data are ordered.

A CLASSIFICATION OF SCALES OF MEASUREMENT

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that

On the Theory of Scales of Measurement. S. S. Stevens. Science, 103 (2684), pp. 677-680, June 1946.

Level of measurement

Level of measurement or scale of measure is a classification that describes the nature of information within the numbers assigned to variables.

The best known classification is that developed by the psychologist **Stanley Smith Stevens**, who proposed four levels, or scales, of measurement: **nominal, ordinal, interval,** and **ratio.**

Stevens proposed his typology in a 1946 Science article titled "*On the theory of scales of measurement"*, in which he claimed that all measurement can be conducted using the four different types of scales, unifying both **qualitative** and **quantitative** data.

Nominal scales are used for labeling variables, without any quantitative value.

"Nominal" scales could simply be called "labels."

Examples include, gender, nationality, ethnicity, language, genre, style, biological species, and form.

Notice that all of these scales are mutually exclusive (no overlap) and none of them have any numerical significance.

With ordinal scales, it is the order of the values is what's important and significant, but the differences between each one is not really known.

Examples include, values such as 'sick' vs. 'healthy' when measuring health, to a spectrum of values, such as Very Happy', Happy', OK', 'Unhappy' when measuring satisfaction.

Here the scale allows for rank order (1st, 2nd, 3rd, etc.) by which data can be sorted, but still does not allow for relative degree of difference between them.

For example, is the difference between "OK" and "Unhappy" the same as the difference between "Very Happy" and "Happy?" We can't say.

Ordinal scales are typically measures of non-numeric concepts like satisfaction, happiness, discomfort, etc.

Interval scales are numeric scales in which we know not only the order, but also the exact differences between the values.

Examples include temperature with the Celsius scale, which has two defined points, separated into 100 intervals, dates, percentages, location in Cartesian coordinates, and direction measured in degrees from true or magnetic north.

Ratios are not allowed since 20 degree C cannot be said to be "twice as hot" as 10 degree C, nor can multiplication/division be carried out between any two dates directly.

However, ratios of differences can be expressed; for example, one difference can be twice another.

The realm of statistical analysis on these data sets opens up because central tendency can be measured by mode, median, or mean; and standard deviation can also be calculated. Ratio scales are the ultimate nirvana when it comes to measurement scales because they tell us about the order, they tell us the exact value between units, AND they also have an absolute zero–which allows for a wide range of both descriptive and inferential statistics to be applied.

Examples include height, weight, duration etc.

These data can be meaningfully added, subtracted, multiplied, divided (ratios).

Central tendency can be measured by mode, median, or mean; measures of dispersion, such as standard deviation and coefficient of variation can also be calculated from ratio scales.

Summary of Data Types

Provides:	Nominal	Ordinal	Interval	Ratio
"Counts," aka "Frequency of Distribution"	~	~	~	~
Mode, Median		~	✓	✓
The "order" of values is known		~	~	~
Can quantify the difference between each value			~	~
Can add or subtract values			~	~
Can multiple and divide values				~
Has "true zero"				~

				Sh	eets Charts Sm	artArt Graphics WordAr
\diamond	Α	В	C	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06
16	70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
17	96	4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

				Sh	eets Charts Sm	artArt Graphics WordAr
\diamond	Α	В	C	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06
16	Item	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
17		4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08
			· · · ·			

				Sh	eets Charts Sm	artArt Graphics WordAr
\diamond	Α	В	С	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06
16	Item	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
17		4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08
-						

				Sh	eets Charts Sm	artArt Graphics WordAr
\diamond	Α	В	С	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	5	2/22/08
4	32	7/16/07	2-High	Small Pack	Attribute 🤅	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06
16	Item	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
17		4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08
-			- · · · ·			

Attribute Types

\circ Nominal

\circ Ordinal

o Interval

\circ Ratio

Attribute Types

\circ Nominal

- o Ordinal
- Quantitative
 Interval
 Ratio

N – Nominal/Categorical (labels) apples, oranges, ...

O – Ordinal movie rating: 1 star to 5 stars

Q – Interval (location of zero arbitrary) Jan 20, 2016; location: (LAT 47 LONG 122) a geometric point cannot compare directly, only differences (i.e. intervals) may be compared

Q – Ratio (zero fixed)

length, mass, ... counts and amounts

				Sh	eets Charts Sm	artArt Graphics WordAr
\diamond	Α	В	С	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06
16	70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
17	96	4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08
-						

				Sh	eets Charts Sm	artArt Graphics WordAr			
\diamond	Α	В	С	S	Т	U			
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date			
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06			
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08			
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07			
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07			
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07			
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07			
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07			
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07			
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07			
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07			
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05			
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05			
14	69	6/4/05	4-Not Specified	Wrap Bag	0.6	6/6/05			
15	70	12/18/06	5-Low	Small Box	0.59	12/23/06			
16	70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06			
17	96	4/17/05	2-High	Small Box	0.55	4/19/05			
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06			
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08			
20	130	5/8/08	2-High	Small Box	0.37	5/9/08			
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08			
22	130	5/8/08	2-High	Small Box	0.6	5/11/08			
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06			
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06			
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08			
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07			
27	166	9/12/07	2-High	Small Box	0.55	9/14/07			
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06			
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08			

				Sh	eets Charts S	martArt Graphics WordAr
\diamond	Α	В	С	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	8 10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.5	5 2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	9 7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.73	2 7/17/07
6	32	7/16/07	2-High	Medium Box	0.0	6 7/18/07
7	32	7/16/07	2-High	Medium Box	0.6	5 7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.5	2 10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.5	B 10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.5	5 11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.4	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.5	6 1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.4	4 6/6/05
14	69	6/4/05	4-Not Specified	Wrap Bag	0.0	6 6/6/05
15	70	12/18/06	5-Low	Small Box	0.5	9 12/23/06
16	70	12/18/06	5-Low	Wrap Bag	0.87	2 12/23/06
17	96	4/17/05	2-High	Small Box	0.5	5 4/19/05
18	97	1/29/06	3-Medium	Small Box	0.3	8 1/30/06
19	129	11/19/08	5-Low	Small Box	0.3	7 11/28/08
20	130	5/8/08	2-High	Small Box	0.3	7 5/9/08
21	130	5/8/08	2-High	Medium Box	0.3	8 5/10/08
22	130	5/8/08	2-High	Small Box	0.0	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.0	6 6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.6	9 6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.87	2 5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	4 10/23/07
27	166	9/12/07	2-High	Small Box	0.5	5 9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.5	7 8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.47	2 4/7/08

			-	Sh	eets Charts Sr	nartArt Graphics WordA
\diamond	Α	В	C	S	Т	U
1	Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
2	3	10/14/06	5-Low	Large Box	0.8	10/21/06
3	6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
4	32	7/16/07	2-High	Small Pack	0.79	7/17/07
5	32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
6	32	7/16/07	2-High	Medium Box	0.6	7/18/07
7	32	7/16/07	2-High	Medium Box	0.65	7/18/07
8	35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
9	35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
10	36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
11	65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
12	66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
13	69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
14	60	6/4/05	4-Not Specified	Wrop Bag	0.6	6/6/05
N	omin	12/18/06	5-Low Ordinal Box		Ouanti	tative 3/06
		12/18/06	5-Low	p Bag	Quanti	3/06
17	96	4/17/05	2-High	Small Box	0.55	4/19/05
18	97	1/29/06	3-Medium	Small Box	0.38	1/30/06
19	129	11/19/08	5-Low	Small Box	0.37	11/28/08
20	130	5/8/08	2-High	Small Box	0.37	5/9/08
21	130	5/8/08	2-High	Medium Box	0.38	5/10/08
22	130	5/8/08	2-High	Small Box	0.6	5/11/08
23	132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
24	132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
25	134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
26	135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
27	166	9/12/07	2-High	Small Box	0.55	9/14/07
28	193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
29	194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

Identify the data attributes

Identify the data attributes

For the popular cricket-playing nations, we took every batsmen that has scored at least 20 runs in their ODI career, and plotted their run rate for every single match. Size = Number of runs. Colour = strike rate 20 50 80 110 140

ODI Test Ind Pak	Aus SL WI	NZ Eng	SA Zi	m W	orld						
Search	Expand Collapse										
Sachin	Mohammad Azharuddin	Mahendra S Dhoni		dra Alaysinhji D Jadeja		Na∨jot S Sidhu		t u	Gautam Gambhir		
R Tendulkar	k	Krishnamacha Srikkanth	Ravish J Sha	ankar astri	S M Ga	unil avaska	ar	Mohai Ka	mmad aif		Virat Kohli
			Vinod G Kamb	Mo Am	hinder Iarnath	Ma M Prat	noj ohaka	Ro G Sh	vhit arma I	lrfan ∢ Pathan	Nayan R Mongia
Sourav	Yuvraj Singh	Kapil Dev	anginuran	A B Ag	ijit Jarkar	Anil Kumble	Yash Shar	ipal ma	lavagal Srinath	Heman <u>o</u> K Badar] Yusuf li K Pathan
C Ganguiy		Vangipurapp S Laxmar		in Dir Mo	nesh ngia	Robin Uthappa	Sadagops Ramesh Praveen	an Roger H Bin Sharma Madan Lai	M Woor ny V Rar Syed M H Kirmani	teri Su nan BJ Saba Karim I	nil Kiran S More
		B Vengsarkar	Rabindr	a Harb	hajan	Raman Lamba	Ashok Malhotra	Nichil Chopra	rtivardhan Sar Azad S D	heer Brijesh ghe PRote	Espu K Velaka Verlaiset Pateruppel Re
Rahul	Virender		R Singr	Krist	nna K	Zaheer Khan	Chetan Sharma	Chandrakant S Pandit	Vilay Seniya Seniya Vilay Vilay Vilay	Raman Cadada R Rowar Cadada Nusil Jashad G Karik Jyub Ki	Jacob Cheandrait Gonan J Martin 9 Chaunan 9 Galador Agan Earolin Karsan Nano Rodal Enginar D Ghavit K Thary
Dravid	Senwag	Suresh K Raina	Sanjay	D Ka	arthik	Ravindra A Jadeja R	Gundappa Viswana t	Praveen Kumar	Jakian Ra Naha Rashore R Bharduaj	banda Buda P anyin Banin Ban Sundarya Anamati Syet L W	An and a series of the series
		i ci tuliiu	∨ Manjreł	(ar San Mi	deep Patil	Pathiv A Patel	Ajay KSharma	Anshuman D Gaekvad	Bangar Rajash Lai K Chauhan Utnder Kannava S Khanna K Rasad	Miay Valendar St	Anne Cing on church an Anne Cing on church an Anne Cing on church an Anne Cing on church an Anne Cing on the Starkard an Anne Cing on the Starkard an Anne Cing on the Starkard an

For the popular cricket-playing nations, we took every batsmen that has scored at least 20 runs in their ODI career, and plotted their run rate for every single match. Size = Number of runs. Colour = strike rate **20 50 80 110 140**

1. In order to visualize data, data sets have to be mapped to visual attributes (also known as data encoding)

- 1. In order to visualize data, data sets have to be mapped to visual attributes (also known as data encoding)
- 2. Understanding data types helps us determine which visual attributes represent which data types most effectively

- 1. In order to visualize data, data sets have to be mapped to visual attributes (also known as data encoding)
- 2. Understanding data types helps us determine which visual attributes represent which data types most effectively
- 3. Classification of data with implications on the mathematical operations we can perform on them and how to structure them

- 1. In order to visualize data, data sets have to be mapped to visual attributes (also known as data encoding)
- 2. Understanding data types helps us determine which visual attributes represent which data types most effectively
- 3. Classification of data with implications on the mathematical operations we can perform on them and how to structure them
- 4. Good data models result in good mental models, leading to better understanding and reasoning

Summary of Data Types

Provides:	Nominal	Ordinal	Interval	Ratio
"Counts," aka "Frequency of Distribution"	~	~	~	~
Mode, Median		~	✓	✓
The "order" of values is known		~	~	~
Can quantify the difference between each value			~	~
Can add or subtract values			~	~
Can multiple and divide values				~
Has "true zero"				~

Distribution of English letters toward beginning, middle and end of words



David Taylor -http://prooffreaderplus.blogspot.com/2014/05/me thodology-and-analysis-of-letter.html

Methodology: www.prooffreader.com/p/explain.html

e.g. the t in "the" contributes more than the t in "thee".

Here are the top 5,000 names of students that appeared in the Class XII State Board examinations in Tamil Nadu. The larger boxes show popular names. The colour of the box indicates the average percentage scored by the student in the Board exams.



REUTERS GRAPHICS

153 studies have been published on the new coronavirus

92 were not peer reviewed

6	First study issued possible link to sn	mentions akes															
The second secon				A Constraint of the second sec		A second se							A distance of the second secon		LTHORAW		A Constant of the second secon
HITHORAN					Angenere Managene			A contraction of the second se		jesosos Beeseas Beeseas	A Design of the second se	A COMPANY AND A					
Sugge				Annual Annua		Landard C	Anna anna anna anna anna anna anna anna		internet. Internetion Internetion	In a second seco	ຄຄ		Language of the second	And			
links to	HIV		An an annual sector of the sec						Annual and a second sec								
			And and a second						A subscription of the subs								Reserves. Martin Status Martin Status
													Mar Harrison			 Antonio Marcine, Santa Santa	

https://graphics.reuters.com/CHINA-HEALTH-RESEARCH/0100B5ES3MG/index.html

Volume I: Russia interference in the 2016 presidential election

10	122		SPIR.	1000		Y days								1																		W	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
		E.			-			-	-	-	#	-	-	-				-		9	Ξ	-	2	Σ	Ŧ	E	1	-		-		4.5	
		-	-		1079 H				1000 1000			1	-	2	2	1.e		-		1900 A		-	-	12	-		1000	2111 21	1				
		靈					W.	1.	1111	1.84							-						101.4	語曲	_	調整	学服		NII I	「「「」			
								-	調売	1.41		104		1			曹		調加					MA	-		10	100 15-			-	-	
			in the second			1000		1		-							(IIII)				Ξ												

Volume II: Obstruction of justice inquiry

- (14)			1000	1				-						-	111		1905 22 -	調整		調節	100	100						商売	Alle.	
	400 300						1			(Tell)			1111	自行					靈云		NIII.						慶二	1	論	
		į.	$\frac{1000}{10\pi}$						Ē	Ê	· 副記 日							憲	11	自力	NII)				備心					<u>AU</u>
		-						曹臣												-			Ē	Ē		檀			100	
		1		180	-		讑					靈						疆	-	書							朣		No.	
	-	-		2000	1												1				靈					- Allah				à

https://www.nytimes.com/interactive/2019/04/19/us/politics/redacted-mueller-report.html