Consulting & Analytics Club
IIT Guwahati

KRACK THE CASE

Guess-It-Matters.

Disclaimer

All information in this guide is accurate and complete to the best of our knowledge. Any data, definitions or inspirations have been acknowledged wherever possible. Any redistribution or reproduction of part or all of the contents in any form is prohibited.

If you find some discrepancy in the material, feel free to contact us at cac@iitg.ac.in

First Published in March 2023
Developed by the Consulting Team at Consulting & Analytics Club, IIT Guwahati.

Explore more resources at iitg.ac.in/sa/caciitg/ktc/

Editorial Team

Managing Editors: Advaith Krishna A (Head of Consulting), Mansi Gunwant (Secretary)

Contributors: Aashrith Ram, Abhishek Singh, Aditya Kumar, Advaith Krishna A, Alina Rizvi, Aman Kumar, Arhita Kundu, Bhadra Tendulkar, Harsh Gupta, Kanike Uday, Mansi Gunwant, Rishi Srinivasan, Rudra Shankar Nandy, Tejas Tendulkar

Preface

"The results are only as good as the assumptions made."

The accuracy of an estimation approach is only as good as the quality of the assumptions made and the methodology used to arrive at the estimate. While the assumptions made play a critical role in determining the accuracy of the estimate, it is also essential to consider the broader methodology used to arrive at the estimate. This includes the data sources used, the statistical techniques employed, and any other methods or techniques used to arrive at the estimate.

Additionally, when it comes to business and strategy, there is seldom an exact answer to any problem or question. It is often left up to the analyst or decision-maker to make educated guesses, also known as guesstimates, in order to provide insights and recommendations. Guesstimates are a crucial component of case interviews, which are often used in the hiring process for consulting and strategy roles. These interviews simulate real-world business problems and require candidates to use critical thinking, problem-solving, and analytical skills to arrive at an estimate or recommendation.

This book is designed to provide a comprehensive collection of guesstimate questions and solutions that can be used to practice for case interviews. The questions in this book have been carefully crafted to cover a wide range of topics and difficulty levels, allowing readers to build and refine their guesstimation skills. Each question is accompanied by a detailed solution that explains the underlying assumptions and approach used to arrive at the estimate, providing valuable insights into how to approach similar problems in the future.

Whether you are a seasoned consultant looking to brush up on your skills or a student preparing for your first case interview, this book is an essential resource for mastering the art of guesstimation. So, without further ado, let's dive into the world of guesstimates and start honing our analytical skills!

- **Advaith Krishna A**
  Mansi Gunwant

Table of Contents

Guesstimates

How should you approach guesstimates?

A general approach towards guesstimates can be thought of as follows:

Clarifying the Scope: Anything that is/may seem ambiguous must be cleared right away. Ask questions! Thoughtful questions show that you have an eye for detail. You need to identify the constraints you are working under to answer clearly and move through the case faster.

Developing a Strategy: Think of an overall strategy that exhaustively covers the issue at hand. Try to think from different perspectives as well. Generally, three methods are used (with tailored variations): Top-Down, Bottom-Up, and Employing a Proxy. Base your strategy on general 2-3 element formulae (such as averages, demand, and supply).

Thinking out loud: At every step of the interview, it is essential to engage with the interviewer. When you have formulated the basic understanding of the problem in your head, start thinking out loud about your strategies and approaches. This will also be beneficial in giving your interviewer an idea of where you might be heading and understanding your approach better, sometimes even in guiding you in case you go wrong.

Break down the problem: Once you identify the constraints you must work under, break down the problem into smaller pieces (a MECE approach). Remember the following factors while breaking them down into parts — Population, Income split, Age split, Gender, and Rural-Urban split.

Perform Math & Sanity Checks: Estimate each piece separately and multiply them to get the overall estimate. Take easy numbers and percentages for your calculation. Perform sanity checks mentally wherever there is "sensitive data" and suggest sanity checks on your own to the interviewer. Making sure your answer is realistic shows that you are methodical in your approach.

This guide follows a slightly different approach, where primary assumptions and a brief overview of the method under consideration are stated under a "Setting the Stage" section, the overall process is discussed under "Approach", and sanity checks are performed under "Validation". Some questions have "Variations" listed at the end for readers who want to practice similar questions.

Tips and Tricks

Averages: The geometric mean is more appropriate for guesstimates than the arithmetic mean when dealing with variables and statistical data, as it considers the compounding effect of the values over time and is less sensitive to outliers or extreme values. This is because extreme values have a disproportionate impact on the arithmetic mean and can distort the estimate of the overall trend.

Fractions: When dealing with large populations or estimating overall trends where demographic variations can be ignored, the fraction of time the average person spends doing some activity is equal to that of people doing it right now. For example, if the average student spends 10% of their time reading, then at this instant, we can expect that 10% of all students are reading.

This… That: Employing a proxy in guesstimates involves using a related metric or data point to estimate the value of another metric or data point that is difficult to measure or estimate directly. A proxy variable can be any quantifiable measure correlated with the variable of interest and is often used when no direct data is available.

For example, if you wanted to estimate the number of cars on a particular road during rush hour but don't have any direct data, you could use the number of people commuting to work in the area as a proxy. You might assume that each person travelling to work drives one car, on average, and use the number of people commuting to work as a rough estimate of the number of cars on the road during rush hour.

Simplicity: Work with simpler numbers, round off intelligently, and use percentages. This is because guesstimates are never about the right answer but about following the best approach. Simplifying numbers makes it easier to work with and perform calculations in your head. In a case interview, time is often limited, and simplifying numbers can help you arrive at an estimate more quickly. This can give you more time to think about other aspects of the problem and come up with a more comprehensive solution.

Supply-side vs Demand-side Approach: Key Insights

Supply and demand are two key concepts in economics that can be used in guesstimates to estimate the potential outcomes of a situation. Both the supply and demand side approaches in guesstimates involve making estimates based on available data, but they focus on different aspects of the problem.

The supply-side approach focuses on the factors that affect the ability of a company or economy to produce and supply goods and services. This approach is often used to evaluate the potential impact of changes in production costs, technology, and the availability of inputs such as labour and raw materials on the ability of a company or economy to produce and supply goods and services.

The demand-side approach, on the other hand, focuses on the factors that affect the desire and ability of consumers to buy goods and services. This approach is often used to evaluate the potential impact of changes in consumer income, preferences, and prices on the demand for goods and services.

The supply-side approach is advantageous when the factors involved in production are well-defined and measurable (e.g. how many cars are fuelled in a gas station). In such cases, it may be possible to make relatively accurate estimates of the potential output based on the available information about the production capacity, resources, and other factors involved.

However, it is also essential to consider the demand side in any guesstimate, even when using a supply-side approach. The demand side provides information about the potential market for a product or service and can help to validate the estimates made on the supply side. For example, suppose a guesstimate suggests that a coffee shop could serve 100,000 customers per year based on its production capacity, but there is only demand for 50,000 customers in the area. In that case, the estimate on the supply side may need to be revised.

Supply-side approaches are easier to calculate, but they can also yield close results, which may not be fully accurate. A rule of thumb is to start from the supply side and perform a sanity check considering the expected demand. In summary, while the supply-side approach helps estimate the potential output of a particular industry, firm or product, it should always be accompanied by consideration of the demand side to ensure that the estimates are valid and reliable.

Some Commonly Used Assumptions for Guesstimates

Rural-Urban Split

World: 60 % Rural and 40 % Urban
India: 70 % Rural and 30% Urban

Income-wise Split in India

Urban: Below Poverty Line - 0%
Lower Income - 40%
Middle Income - 40%
Upper Income - 20%

Rural: BPL- 20%
Lower Income - 30%
Middle Income - 40%
Upper Income -10%

Age Split

0-14: 25%
15-25: 20%
26-59: 40%
60+: 15%

Gender Split

Female: 50%, **Male:** 50%

Occupation-wise Split

Global: People of Working Age - 50% of the total population
of which, Blue Collar - 75%
White Collar - 25%

India: People of Working Age - 50% of the total population
of which, Blue Collar - 75%
White Collar - 25%
(or)
Agriculture and Fishery - 40%
Industry and Allied Sectors - 25%
Services - 35%
(or)
Unorganised Sector - 70%
Organised Sector - 30%

Internet Penetration in India: 40%

Population

World: 8 billion
India: 1.4 billion or 140 Cr
USA: 330 million or 33 Cr
UK: 68 million or 6.8 Cr
Canada: 4 Cr
Delhi: 3 Cr
Bangalore: 1 Cr

Household size

USA: 2.5
India: 4.4 (~ 4 in urban and ~ 5 in rural areas)

Guesstimates

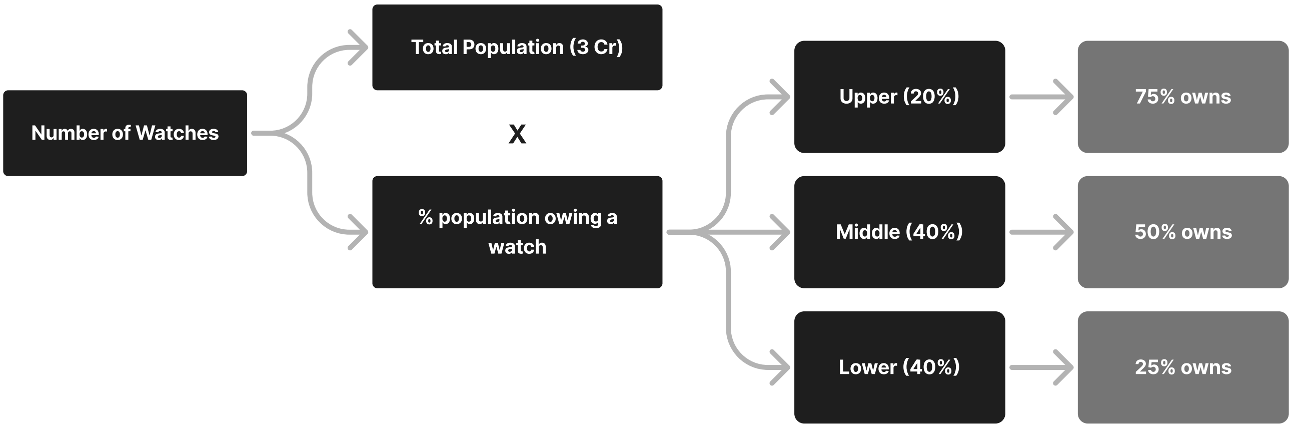
Estimate the number of watches currently in use in Delhi.

Setting the Stage

This question can be approached considering an income, household, or age-wise split. The following approach is **based on the income-wise break** of the population. The rural population of Delhi (around 2% of 3 Cr, the total population) can be ignored for simplicity. We'll assume that one would own up to one watch at any time. Since watches are easily replaceable/repairable, we can consider that all are in working condition.

Approach

The approach can be detailed as follows:



This would give us a total of **3 Cr** $\times \left(\frac{3}{20} + \frac{4}{20} + \frac{2}{20}\right)$ **= 1.35 Cr watches.**

Variations

Another problem uses a similar approach: **Estimate the number of laptops in working condition in Delhi.**

Here, you'll have to divide this problem into two parts — personal and work laptops. For personal laptops, you might need to consider that lower sections will have one laptop per household, while the upper-income population might have two per household. Additionally, we can assume that there are four members in every household. 25% of the working population (1.5 Cr) will use a PC/laptop at their office. It is safe to assume that a quarter of this would use a laptop.

The number of personal laptops = ($\frac{3}{4}$ **Cr households)** $\times \left(2 \times \frac{3}{20} + \frac{4}{20} + \frac{2}{20}\right)$ **= 45 lakhs.**

The number of work laptops = 1.5 Cr × 0.25 × 0.25 = 9.375 lakhs ~ 10 lakhs. Assuming 10% of all laptops don't work at any point, we'll get ~ **49.5 lakh laptops**.

Estimate the number of aeroplanes flying in India right now.

Setting the Stage

The number of aeroplanes currently flying would include those **that have taken off but have yet to land**. In this case, we'll consider only commercial passenger flights while **ignoring any rush caused by seasonal demand**.

One method to approach a problem like this is to find the average flight duration and then multiply that value by the number of flights taking off in that period. We'll follow a **supply-side approach** for convenience since segmenting the airports would be easier based on the gap between take-offs. Since the number of aeroplanes taking off would vary depending on the time of day and the type of airport, we'll **assume total operational capacity and later adjust it to arrive at a range**.

Approach

We can assume the average flight duration to be around 2 hours (a quick way to verify this is to take a known range and find the geometric mean — in this case, it should be between 1 and 4 hours hence giving an average of 2 hours).

Now, we can split the airports into three types based on capacity and estimate the number of flights taking off in 2 hours. We'll consider 130 airports in total.

Type	No. of Airports	No. of Runways	Gap between take-offs (in min)	No. of flights taking-off in an hour
Major	5	2	5	120
Medium	25	1	10	150
Small	100	1	30	200

The total number of flights taking off in an hour = 470, and the total number of flights taking off in two hours = 940 ~1000. Assuming that the airports operate at 75% of their total capacity during peak hours and at 25% during the least active hours, the number of aeroplanes flying anytime would be within a range of 250-750 (taking the geometric mean of this range will give us **~400 flights in the air at a given time**)

Validation

We will find the number of people flying daily to verify whether this number is correct. Assuming an average aeroplane has 150 seats with 80% occupancy, this would amount to 120 occupied seats per flight. Now 400 flights will carry 48,000 passengers. Considering most travel happens in 16 hours (or eight slots), there would be close to 4 lakh passengers travelling daily, which seems reasonable.

How many cups of tea are consumed in all of India every week?

Setting the Stage

We'll assume that every day of the week is considered identical and does not affect tea consumption and that the number of people who take tea more than twice a day is negligible. Consumption varies depending on the age group; hence we'll **use an age-wise split** to estimate the weekly consumption for the whole population.

Approach

The total population of India is 140 Cr. We've assumed that most people in the age group 0-14 would prefer milk or coffee, which is evenly distributed in the 15-25 group. Tea is typical in older age groups. Hence, we can find the average daily consumption per person as follows:

Age group	(A) % of people of that age group	(B) No of cups per day per day	(C) Further division	Daily Consumption (A × B × C)
0-14	25%	0	90%	0
		1	10%	0.025
15-25	20%	0	40%	0
		1	50%	0.1
		2	10%	0.04
25+	55%	0	10%	0
		1	50%	0.275
		2	40%	0.44
			Daily Consumption per person	0.88
			Weekly Consumption in Cr (0.88 × 7 × 140)	**862.4**

Hence, ~ 850 Cr cups of tea are consumed in India weekly.

Variations

The following questions can be easily solved using an age-wise split. Do try them out!

1. How many video games are sold across all platforms in India annually?
2. How many cricket bats are sold in India annually?
3. What is the average age of people who buy cars in India? (This one's a bit tricky, state your assumptions clearly before you proceed)

Estimate the size of the chewing gum market in India.

Setting the Stage

We'll assume that the number of people consuming chewing gum in the age group below 15 and above 50 is negligible and can be ignored. Also, we **ignore the consumption by the lower-income population** (50% of the total population). Consumption varies depending on the age group; hence we'll use an age-wise split to find the daily consumption per person and estimate the yearly consumption for the population.

Approach

The total population of India is 140 Cr. Chewing Gums are typically consumed by middle and higher-income people (50% of the total population). Thus the population in focus is 70 Cr. We categorise the consumers as high-frequency, medium-frequency and low-frequency consumers. Also, we assume the price of one packet, which contains four units of gum, to be Rs. 5, and the number of weeks in a year to be 50. Hence, we can find the market size of chewing gums as **No. of Consumers × No. of packets per year × Price per unit.**

Age group	No. of people in that age group (Cr)	(A) No.of packets per week	% consumers	No. of Consumers (Cr)	No.of packets per year(A* 50)	Yearly Consumption
0-14 (30%)	20	0	-	-	-	0
15-17 (Low, 10%)	7	2	0	0	100	0
		1	20%	1.5	50	375
		0.5	80%	5.5	25	687.5
18-24 (High, 10%)	7	2	60%	4	100	2000
		1	30%	2	50	500
		0.5	10%	1	25	125
25-35 (Medium, 15%)	11	2	5%	0.5	100	250
		1	60%	6.5	50	1625
		0.5	35%	4	25	500
35-50 (Low, 15%)	11	2	0%	0	100	0
		1	40%	4.5	50	1125
		0.5	60%	6.5	25	812.5
50+ (20%)	14	0	-	-	-	0
Market Size (in Cr)						8000

Hence, the market size of chewing gums in India is ~ 8000 Cr.

Variations: More Market Sizing Questions

The following questions are based on finding the annual market size of different products and will follow a similar approach. Do try them!

1. Estimate the market size of formal ties in India.
2. Estimate the annual size of socks sales (by revenue) in India.
3. Estimate the size of the diaper market in India.

Estimate the number of ATMs in Delhi.

Setting the Stage

We'll find the average number of ATM cards in a household, and the **distribution of the population of Delhi based on population density** and find the number of ATMs required.

Approach

The population of Delhi is 3 Cr. Assuming 4 people per household, there should be 75 lakh households in Delhi. Now, we'll estimate the total number of ATM cards in Delhi as follows:

Income Level	% of population	ATM Cards per household
Lower	40	1
Middle	40	2
Upper	20	4

The average number of ATM cars in a household = 2. Therefore, there are 1.5 Cr ATM Cards in Delhi. Now, 60% of the population resides in high-density areas, and the rest in low-density areas. Let's assume we have 1 ATM per 5000 cards in high-density areas while 1 ATM per 2500 ATM cards in low-density areas.

This gives a total of **(90 lakh / 5000 + 60 lakh / 2500) = 1800 + 2400 = 4200 ATMs**.

Validation

For an ATM in a high-density area, an ATM operating for 16 hours daily can process 12 × 60/3 = 240 transactions in a day (assuming 3 minutes are taken for a transaction) or 7200 transactions per month. Since 5000 cards use one ATM, this would mean **that one card is used at least once every month.**

How many packets of noodles are consumed monthly at a college campus?

Setting the Stage

Let's assume the average number of students in a residential college to be around 7000. Each department, on average, has 50 professors, which means 400 faculty members across eight departments (say). The number of other technical and administrative staff = 75% of academic personnel = 300. This means the total number of working professionals = 400 + 300 = 700. We should also include their family members, an average of one per member. We assume that other working professionals, like hostel workers, guards etc., do not reside on campus; thus, their number would not contribute to the total value. Additionally, we'll consider that there are four weeks per month.

Now that we've some data to work with, we'll estimate the consumption for students and working professionals (along with their families).

Approach

Consumption by Students

Frequent - Students eat an average of 4 packets of noodles weekly. This would be 20% of the student population which implies 1400 students. This implies 5600 packets of noodles per week and 1400 × 4 × 4 = 22,400 packets in a month.

Less Frequent - Students that eat two packets of noodles every week. This would be roughly 50% of the student population for a college, which comes out to be 3500 students. This implies 3500 × 2 × 4= 28,000 noodles in a month.

Least Frequent - Students that eat one packet of noodles every week. This is the rest of the student population, which means 2100 students and (2100 × 1 × 4) = 8400 packets per month.

Therefore the total number of **packets consumed by students would be 58,800**.

Consumption by Professionals

For working professionals, we can again divide the population into three types. However, the number of packets consumed per person would be lower compared to students, given that college students are more likely to eat noodles.

Frequent - Families where every person eats three packets of noodles every two weeks, which means six packets of noodles every month per family member. The number of such families will again be 20%. This gives a number of (0.2 × 700 × 2 × 6) = 1680 packets per month.

Less Frequent - Families where every person eats two packets of noodles every two weeks, which means four packets of noodles every month per family member. The percentage of such families will be 50%. This results in (0.5 × 700 × 2 × 4) = 2800

Least Frequent - Families where every person eats one packet of noodles every two weeks, which means two packets of noodles every month per family member. The number of such families will be 30%. This gives (0.3 × 700 × 2 × 2) = 840

Therefore the total number of **packets consumed by professionals would be 5320**.

Total number of packets = 64,120 ~ 65,000 per month

Validation

The total number of people considered for the guesstimate is 7000 + 1400 = 8400
The number of packets consumed in a day = 65,000 / 30 = 2166.67
Number of packets consumed by one individual in one day = 2166.67 / 8400 = 0.26
This means that, on average, a person in the college consumes roughly one packet of noodles every four days, which is a reasonably acceptable rate.

How many hotel-sized mini bottles of shampoo are sold every year?

Setting the Stage

We can find the number of bottles produced from a supply-side approach, factoring in the average consumption and occupancy.

Approach

We'll assume that there are 4000 major cities worldwide, and each such city will have around 15 hotels, of which 10 provide shampoo bottles to every customer. Additionally, each hotel will have around 50 rooms, and most customers will stay for two days. Consider that for 25% of the year, a hotel remains 100% occupied and 50% occupied for the rest of the year, so the average occupancy is around 62.5%. This means (4000 × 10 × 0.625 × 50) = 12,50,000 rooms are occupied daily worldwide.

We'll divide the rooms in a hotel into three categories :

A. **Studio Rooms:** 30% - 1 bottle in a two-day stay
B. **Suite rooms:** 40% - 1 bottle in a two-day stay
C. **Family rooms:** 30% - 2 bottles in a two-day stay

So we have 70% of occupants using half a bottle per day and 30% of occupants using one bottle a day, which gives **((0.7 × 0.5) + (0.3 × 1)) × 12,50,000 = 8,12,500 bottles per day or 8,12,500 × 365 = 29,65,62,500 ~ 30 Cr.**

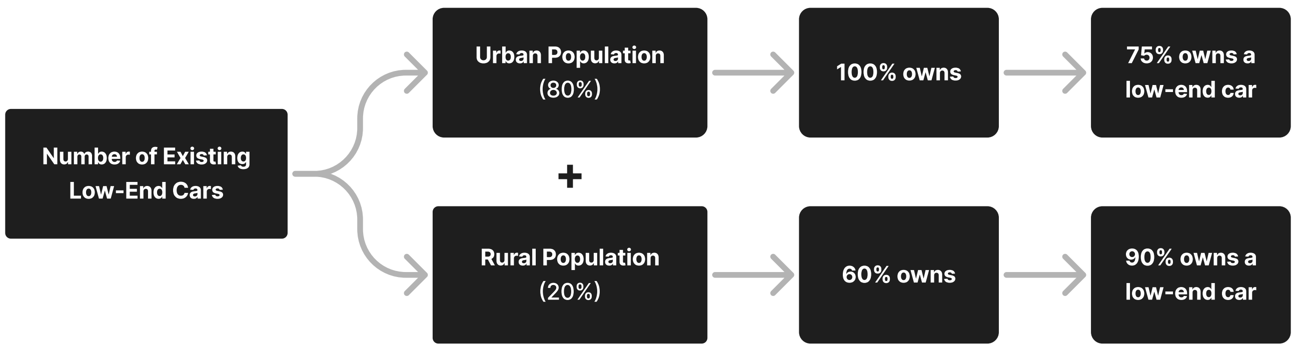
How many Mazda dealers are there in the USA?

Setting the Stage

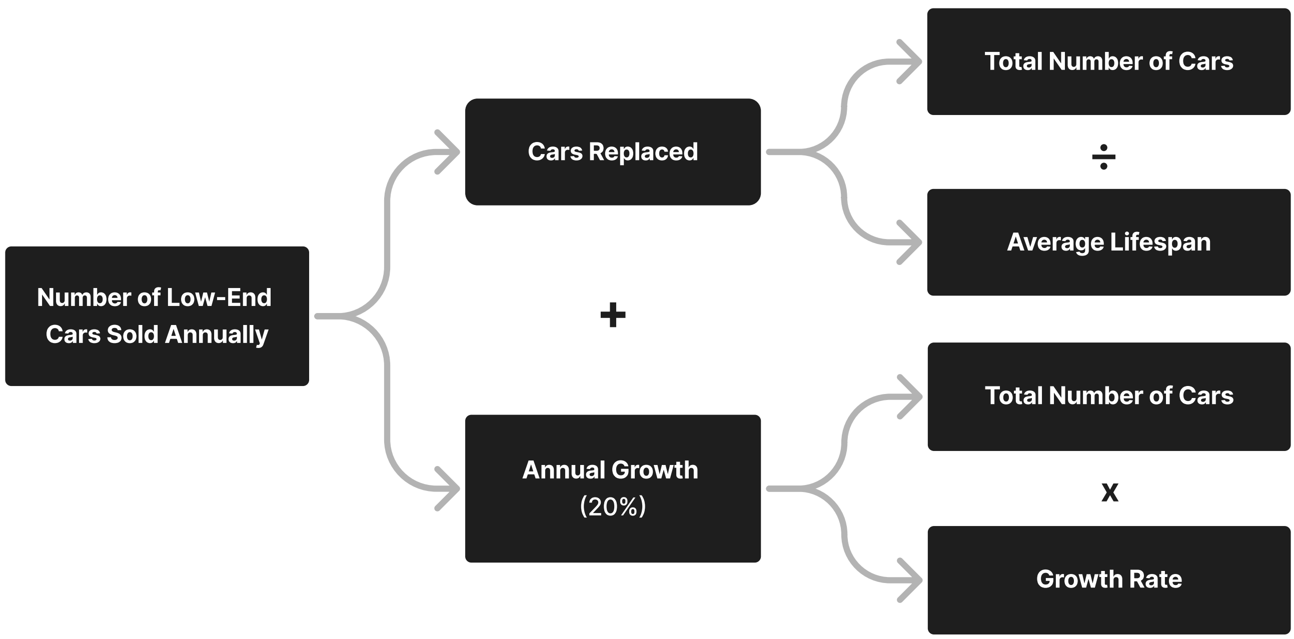
Mazda is a minor player in the US (~2% market share) and is usually cheap. Their cars would fall in the $15,000 - 30,000 bracket. Additionally, we'll consider that the average household size in the US is 2.5 (compared to 4.4 in India). Further, the **average lifespan of a car is taken as 6 years.**

Approach

The number of households in the US = 33 Cr/2.5 = 13.2 Cr households.



The number of existing low-end cars = **13.2 Cr × (0.8 × 1 × 0.75 + 0.2 × 0.6 × 0.9) = 9.34 Cr.** The number of low-end cars sold annually can be found as follows:



We'll assume that the population grows at 1%. This gives us a **9.34 × 0.01 + 9.34/6 = 1.65 Cr low-end cars.** Now out of this, Mazda cars would amount to ~ 3.3 lakhs. Assuming a dealer sells 50 cars every month or 600 cars yearly, this would amount to 550 dealers in the US.

How much change would you find on the floor of an average mall in a day?

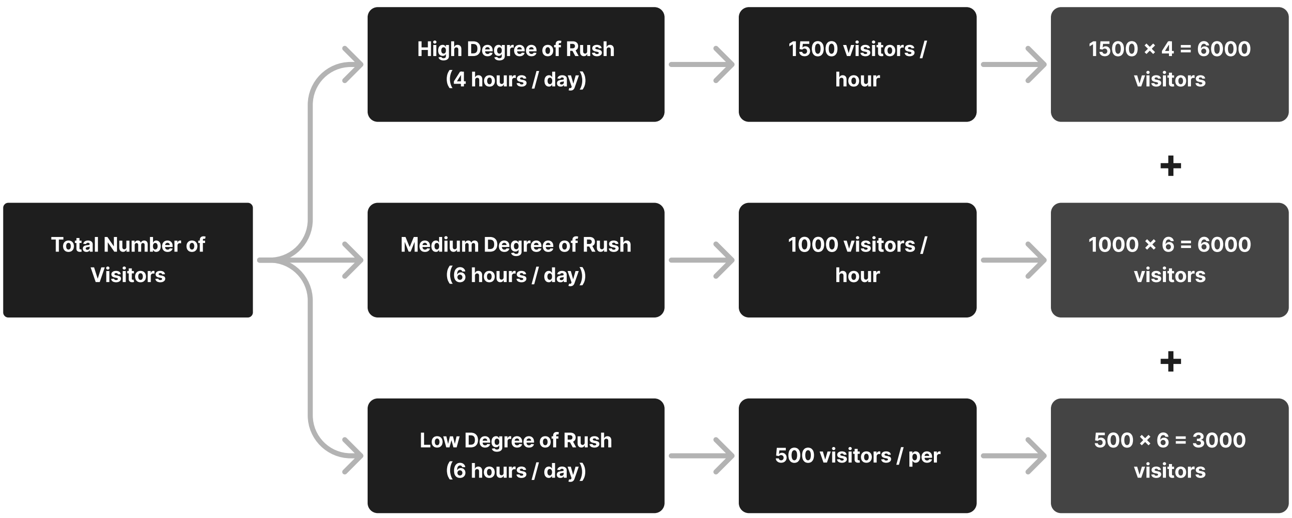
Setting the Stage

In this case, we'll assume that change is only found as coins (and not cash). We'll consider the mall remaining open for 16 hours daily. With a large population, it is safe to assume that each customer takes the same amount of time to complete a purchase and that each customer only visits one store.

Approach

To find the number of stores in a mall, we'll find the total usable area by the space occupied by an average store. An average mall will have two floors, each occupying around 400,000 sq ft with a **usable floor area of 300,000 sq ft.** This gives us a total usable area of 600,000 sq ft. An average store will occupy a space of 100 ft $\times$ 100 ft = 10,000 sq ft, meaning 600,000/10,000 = **60 stores in an average mall.**

A mall would experience peak rush during 4 hours in the night (say between 6 and 10 PM) and the least occupancy close to opening (morning) and closing hours (late night). During peak hours, an average store (having two counters) can complete billing of a single order within 5 minutes, giving us a total of **60 $\times$ 12 $\times$ 2 = 1440 ~ 1500 people visiting the mall during a 'peak' hour.**



This would give us **6000 + 6000 + 3000 = 15,000 visitors daily.** We can assume that 50% of the customers would pay by cash, and 5% of every customer paying by cash will lose a coin. Coins have a value of 1, 2, 5 or 10 rupees — then an average person is likely to lose 2 rupees. This would give us around **2 $\times$ 0.05 $\times$ 0.5 $\times$ 15,000 = 750 rupees.**

Validation

Note that **although we've ignored the effect of anchor stores in a mall, this will be compensated by the number of counters in a store under consideration.**

How many words are there in an average daily newspaper?

Setting the Stage

Let's assume the **dimensions of a newspaper to be 12 in × 20 in with a total of 24 pages**. The idea is to **find the usable space** (occupied by text) by removing space occupied by margins, illustrations and borders from the total area of a newspaper. Further assumptions are listed as follows:

1. A standard page has about 33% illustrations.
2. The thickness of borders = 4 inches on top, 2 inches on other sides
3. Number of pages with 90% or more area covered in illustrations = 4 (i.e. negligible amount of text)
4. Average word density per square inch = 30 for standard text and 8 for headings

Approach

The total area of a page = 240 sq in
The space occupied by images = 80 sq in
The space occupied by borders = (4 × 12) + (0.5 × (9 + 16 + 16)) = 56.5 sq in

The area remaining = 103.5 sq in

Accounting for spacing between columns, the area remaining comes down to somewhere around 100 sq in — the net space filled with text. 90% of this would be occupied by standard text and 10% by headings.

Number of words per page = **(90 sq in × 30 words per sq in) + (10 sq in × 8 words per sq inch) = 2780 words**

Number of words in a newspaper = **2780 × 20 = 55600 (roughly around 50,000)**

Validation

An average article in a newspaper would have upwards of 250 words. This should mean around 200 articles in a newspaper, which seems reasonable.

Variations

Try this question: **Estimate the number of unique words in a newspaper.**

Here, you can write a couple of sentences to find the percentage of unique words. Notice a pattern of repeating words? Try extrapolating this to sentences of median length. Once you find the total number of words in a newspaper, you can multiply this percentage to see the number of unique words. Ideally, your answer should come somewhere around 2000 - 6000 words, depending on the quality of writing in a newspaper.

Estimate the number of flights taking off in a day from Delhi Airport.

Setting the Stage

In this case, we'll consider only commercial passenger flights while ignoring any rush caused by seasonal demand.

One method to approach a problem like this is to consider the total number of boarding gates at the Delhi Airport and calculate the capacity of these boarding gates at different hours of the day. We'll follow a **supply-side approach** for convenience. Since the number of aeroplanes taking off would vary depending on the time of day, we'll divide the hours into peak, non-peak and non-operational.

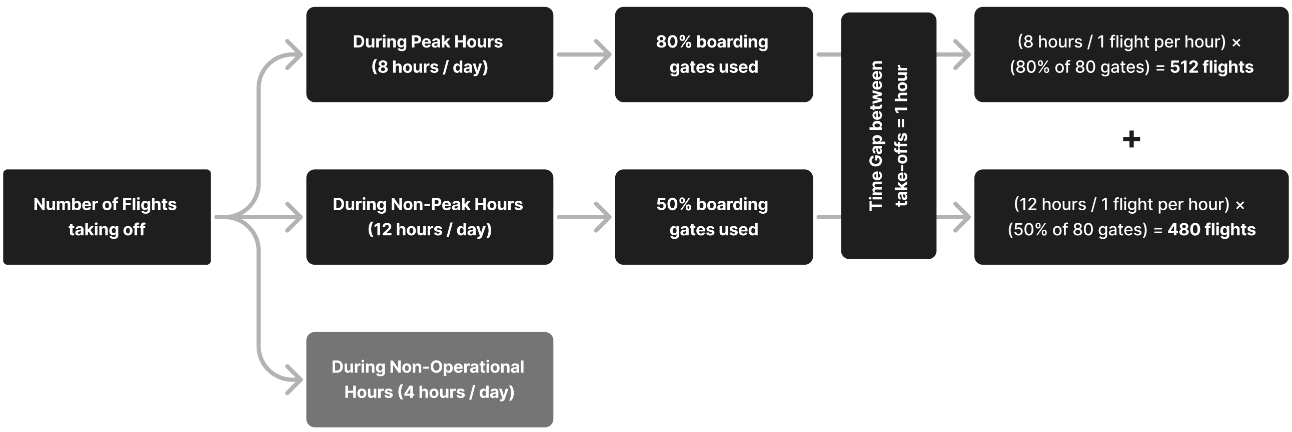
Approach

We can assume the number of boarding gates at Delhi Airport to be 90. The time taken by one flight to take off, which includes boarding time, is around 1 hour.

The hours are divided into peak, non-peak and non-operational. Assume that the airports operate at 80% of their total capacity during peak hours, 60% during non-peak hours, and unoccupied at non-operational hours.

Formula: (No. of boarding gates × Peak hours × %boarding gates occupied at that hour) / Time gap between take-offs

Total number of boarding gates = 80



Therefore, the total number of flights taking off in a day from Delhi Airport is **512 + 480 ~ 1000.**

Validation

We will find the number of people flying daily to verify whether this number is correct. Assuming an average aeroplane has 150 seats with 80% occupancy, this would amount to 120 occupied seats per flight. 945 Flights would carry approximately 1 Lakh Passengers, which seems like a reasonable amount of passengers travelling from the Delhi Airport per day.

[Practice] Estimate the daily revenue of an average fuel station in Mumbai.

This question is similar to the previous example. One can start with identifying various revenue streams — an average fuel station would have four booths with two outlets each. Most fuel stations sell other petroleum-based products and offer tyre-pressure checks. The revenue earned through the latter streams can be ignored as they are negligible. Once you figure out how many cars visit a fuel station daily, you can determine how much each customer would spend on refuelling their vehicle (consider two- and four-wheelers separately since the average billing amount would differ in each case).

Once you've arrived at an answer, you can perform a sanity check by looking at how frequently one refuels their vehicle.

[Practice] How many fuel stations are there in India?

You can start by looking at the number of vehicles in India. Now we need to estimate the total demand for a day/week/year. Once you have an idea about the total demand for fuel stations, find the demand for one fuel station. You have probably found this in the previous practice question — this is the number of cars visiting the fuel station in that particular timeframe. Dividing total demand by the local demand will give you the answer.

Estimate the number of customers purchasing smartphones from Flipkart during the Big Billion Day Diwali Sale.

Setting the Stage

In a problem like this, it is necessary to factor in the customer demographics, such as their income status, internet penetration in their region, and the locality of their stay. This will affect the likelihood of their online shopping, thereby, our estimation for this particular question. Additionally, we must take into account the traffic that these online sites experience during seasons like Diwali.

Approach

The population of India is 140 Cr. or 1.4 Bn. Flipkart will be accessible to areas where people will have access to the internet. Internet penetration is 40%. So the population in focus is 1.4B × 40%. The distribution of the population in urban and rural areas is 30:70. We will divide each of the rural and urban segments based on income split and then calculate the average percentage of people likely to purchase smartphones online.

Population in focus= 1.4Bn × 0.40 = ~ 0.6 Bn

Locality Split		Income Split		(B) % likely to purchase online	Number Likely to purchase(Billion)
Urban	30%	Lower Income	40%	30%	0.021
		Middle Income	40%	70%	0.050
		Higher Income	20%	80%	0.028
Rural	70%	BPL	20%	0%	0
		Lower Income	40%	20%	0.033
		Middle Income	30%	30%	0.037
		Higher Income	10%	50%	0.021
				Total (In Bn)	**0.19**

Hence, the **total number of customers likely to purchase smartphones online is ~ 0.2 Bn**.

The market share of Flipkart, Amazon, and other online electronics stores like Croma or Reliance Digital is 40 %, 40%, and 20%, respectively. Therefore 40% of the customers likely to purchase smartphones online will purchase smartphones from Flipkart. This number is (0.40 × 0.2) Bn = 0.08 Bn

An average person replaces their smartphone once in three years; therefore, the number of people buying in a particular year is 0.027 Bn. The sales during the Diwali season for Flipkart are expected to be more than the average sales made during a month. Assuming that about 50% of the sales come from the Diwali season, the number of people buying smartphones from Flipkart during the Big Billion Day sale in a particular year = **0.50 × 0.027 = 0.013 Bn or 1.3 Crore.**

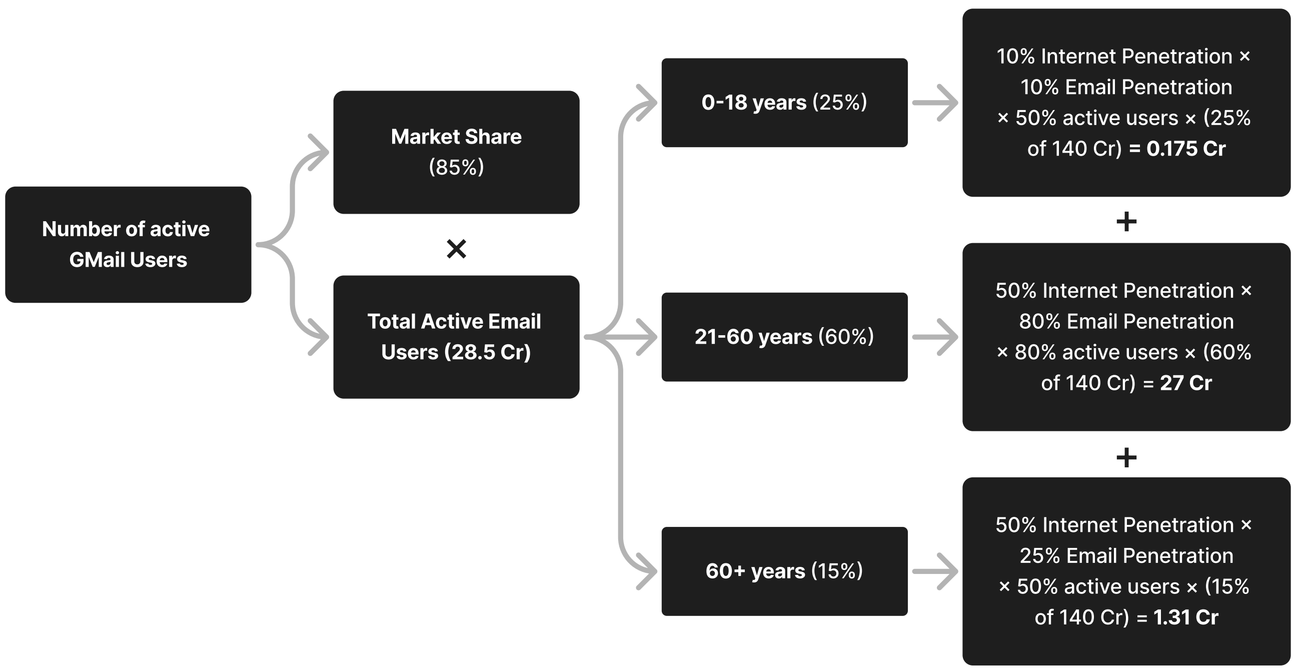
Determine the number of active GMail users in India.

Setting the Stage

For this problem, we need to factor in the internet penetration in India, which is 40%. We also consider an age-wise split and then divide the users on the basis of their penetration of usage and need for an e-mail account. You can also segment the users based on occupation and go ahead.

Approach

The population of India is 140 Cr. Now, we divide the users on the basis of their ages and then calculate their internet penetration and penetration of email usage. Finally, we determine the monthly active users and take the market share of GMail to 85%.



Therefore, the total number of active users in India of GMail is ~ **24 Cr.**

Variations

Try this one out: Estimate the number of daily users on WhatsApp. What if you were asked to find the total number of messages sent daily?

How many medium-sized factories are there in Mumbai?

Setting the Stage

The definition of a medium-sized factory can vary depending on the sector and region. Generally, it is based on the number of employees, annual revenue, or other factors. We'll define it on the basis of blue-collar workers employed at the factory. We can now find the working population

Approach

Consider that a medium-sized factory would employ a minimum of 50 and at most 200 blue-collar workers. We'll take the geometric mean and estimate 100 blue-collar workers. Now, Mumbai has a population of 2 Cr, of which 1 Cr form the working population and 75 lakh people do blue-collar jobs. Assuming 10% work in factories and 30% of factory workers work in medium-sized factories, we'll have 2.25 lakh people employed in medium-sized factories. This means there should be 2,25,000/100 = 2250 medium-sized factories in Mumbai.

Validation

Mumbai is one of the major commercial and industrial hubs in India and has a diverse range of manufacturing industries. While the exact number of medium-sized factories in Mumbai may vary depending on the definition and criteria used, 2250 seems like a reasonable estimate given the city's size, economic profile, and industrial activities.

How much money is collected per hour during the rush hour at a toll plaza on the Delhi-Mumbai Expressway?

Setting the Stage

Let's assume the average toll amount to be INR 50 and the number of lanes to be 8. Now, we can estimate the total amount collected in an hour from the time taken per car per lane and the average toll amount. We'll assume the toll plaza operates at peak capacity during rush hour.

Approach

The toll booths in India operate cashlessly and will take around 8 seconds per car per lane. This means one lane can handle 450 cars in an hour, and the toll plaza can handle 8 × 450 = 3600 cars per hour.

Then the amount collected will be 3600 × 50 = INR 1,80,000.

Estimate the size of the diaper market in India.

Setting the Stage

For this problem, we're taking into account diapers used by babies, as well as elderly people, and those with diseases. We will adopt a demand-side approach. The key factors that will affect our calculation are the total number of people in the age group likely to use diapers, daily diaper demand, and the price of each diaper.

Approach

The population of India is 140 Cr. or 1.4 B. We will divide our consumers into two parts: Infants and Adults. We will take the price of each diaper to be Rs. 10. We also assume that the average number of diapers used daily by infants is 5 and that by the elderly is 2. To approximate the number of people in the 0-3 age group, we take the life expectancy to be 70 and uniform age distribution among all ages. Therefore, the number of people in the age group 0 to 3 years = (Population of India / Life Expectancy) × Age Range= (140/70) × 3 = 6 Cr.

Now, for infants, Total Diaper Market Size = Number of people who can afford × Daily Diaper Usage × 365 × Price of each diaper.

Consumer	No. of Infants (Cr)	Income split	% people in income range	% people who can afford	Number of people (Cr) who can afford.	Total Diaper Market Size (Cr)
Infants	6	Higher Income	10%	100%	0.6	10950
		Middle Income	30%	70%	1.26	22995
		Lower Income	60%	10%	0.36	6570
Total						40515

For adults, Total Diaper Market Size = Number of people who need diapers × No. of diapers used per day × 365 × Price of each diaper.

Consumer	Age	% people in age range	No. of People (Cr)	% people who need diapers	Number of people (Cr) who need diapers	No. of diapers used per day	Total Diaper Market Size (Cr)
Adults	3-25	-	-	0	-		
	26-59	40%	56	2%	1	2	7300

Consumer	Age	% people in age range	No. of People (Cr)	% people who need diapers	Number of people (Cr) who need diapers	No. of diapers used per day	Total Diaper Market Size (Cr)
	60+	15%	42	10%	4.2	2	30660
Total							37960

Therefore, the total market size of diapers in India is approximately 400 Bn + 380 Bn = 780 Bn.

Validation

Let's find out the monthly spending on diapers. Assuming an average household size of 4, the average monthly spend is ~ INR 180. However, not all households have children or adults requiring diapers, which has resulted in lower spending than expected. If only 25% of households had the target audience, the monthly expenditure would be ~ INR 720.

[Practice] Estimate the size of the Indian Cement Industry.

Start with the number of houses and other buildings based on replacement rates and population growth. With this data, you can determine how many walls an average building has and its thickness. Remember to subtract the space consumed by doors, windows and vents. Once you find the price per unit (volume), you can find the market size.

[Practice] Estimate the number of traffic signals in your city.

Start with the number of roads running North to South and East to West. Find the number of intersections and the percentage of roads that might have traffic signals. Note that your answer is highly sensitive to assumptions (as roads might run in other directions also). You'll probably arrive at an answer that might not be even close, but note that these types of questions are only meant to test your approach. Another approach is to find the total area of a city, the area of a block and the number of lights per block. Also, try this one out: Estimate the total length of roads in the US.

Estimate the number of cigarettes smoked in India per month.

Setting the Stage

For this problem, we'll consider multi-level splits, including age, locality, and gender. All of these factors will significantly affect the consumption of cigarettes. We will ignore the age group below 16 for this question as they consume negligible amounts of cigarettes.

Approach

The population of Hyderabad is 0.9 Cr or 9 Mn. Key Considerations for our question are that due to differences in buying power, the urban population consumes more no. of cigarettes than the rural population. In general, females tend to smoke lesser than men.

Consumer	Locality Split	Gender Split	Penetration of People smoking	(A) Population that smokes (Cr)	(B) Number ofc(igarettes per day	(A × B) Total cigarettes per month
<16 Years(30%)	-	-			0	-
16-22 Years(10%)	Urban(40%)	Male(50%)	50%	1.4	60	84
		Female(50%)	10%	0.28	30	8.4
	Rural(60%)	Male(60%)	25%	1.26	30	37.8
		Female(40%)	5%	0.168	10	1.68
22+ Years(60%)	Urban(30%)	Male(50%)	60%	7.56	90	680.4
		Female(50%)	20%	2.52	60	151.2
	Rural(70%)	Male(60%)	20%	7.056	40	282.24
		Female(40%)	5%	1.176	20	23.52
Total						1269.24

Therefore the number of cigarettes consumed in India per month is 1269 Cr.

[Practice] Miscellaneous

How much will my monthly phone bill amount to in the next five years?

In a problem like this, one should ask the interviewer for more information about the category of the user being referred to — their current phone plan and usage, such as what type of plan they have (e.g., unlimited data, prepaid or postpaid, family plan), what their monthly bill is currently, and how many lines they have on their plan. Identify the key factors that will impact the phone bill over the next five years: data consumption and requirements changes, inflationary changes, etc. The idea for such a question is not to give an exact answer but to communicate your approach to the interviewer.

Editor's Note
A variation of this question was asked in a KPMG interview with the following considerations: assume that the user has a prepaid plan offering unlimited calls and 1.5 GB of daily data. The user has opted for a family billing plan with their spouse (enrolled in the same plan) and their 10-year-old child (registered on a plan which costs half theirs, offering unlimited calls and 6 GB monthly data).

How many penguins can live in Antarctica?

The question sure looks funny, but variations have been found to be asked in multiple interviews. For a question like this, ask for more information about the size and geography of Antarctica, as this will give you a better understanding of the potential habitat available for penguins. With this information, you can estimate the carrying capacity of Antarctica for penguins. This will depend on factors such as the amount of available habitat, the availability of food, and the presence of predators or competitors.

Editor's Note
What if the interviewer asked you to factor in the effects of climate change? How will this change your answer?

Are there two dogs in the world with the same number of hairs?

Take inspiration from the pigeon-hole principle and the birthday problem. The probability is $p(n) = 1 - \frac{h!}{h^n(h-n)}$ if $n \leq h$ and 1 otherwise for 'n' number of dogs and the maximum combination of hairs 'h'. Now, if you believe that there are over 1 million dogs in the world and less than 1 million hairs on a dog, the odds are 100% that at least two dogs in the world have the exact same amount of hair.

Guess-It-Matters.

Consulting & Analytics Club, IIT Guwahati
iitg.ac.in/sa/caciitg