Guesstimates play an important part in any consulting or analytics interview.
They are estimates based on a mixture of guesswork and calculation. The process of problem-solving is fairly simple. Look at the parameters that might affect the given problem and arrive at the estimated quantity. An example could be finding the number of red lights in Delhi.

General Framework

1. Understand the question (Clarify) – Try and extract the exact information from the interviewer on what he wants you to calculate. In the previous example, for all you know, the interviewer might be interested in finding out the total number of red light bulbs in Delhi. Hence, it is imperative to be on the same page as the interviewer. Always ask first before you attempt any guesstimate.

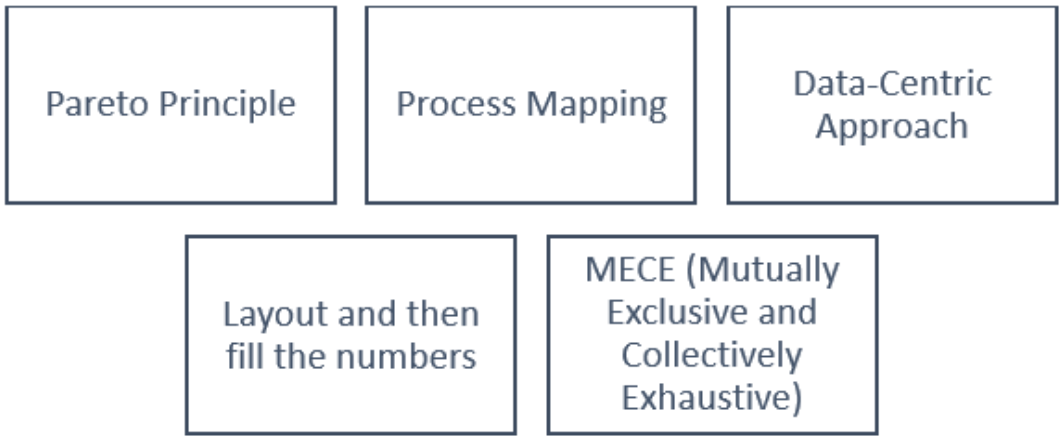
2. Devise a Logical Approach (Structure) – There is no foolproof way to approach a guesstimate. You can solve a guesstimate using several approaches – top-down, bottom-up, process mapping, layout centric, or critical comparison. The trick here is to go with the approach that helps you minimize your assumptions, a simple rule of thumb while you approach guesstimates.

3. Decide which Approach to take (Analyze) – Once you have devised the approaches, it is time to decide which approach you want to use. Ideally, think of 2-3 steps ahead of you in the approach you decide to use and see how it works.

4. Put the numbers and the assumptions (Conclude) – This is the final and the trickiest step. Once you get this right, you have cracked the guesstimate. However, once you practice enough, you develop a knack for solving it.

5. Do a sanity check – itis a method to check if the answer from the guesstimate is in the bounds of reasonableness

Strategies to solve Guesstimate:



Let us look at some techniques to solve guesstimates:

1. Pareto Principle (80:20 Rule): The idea behind this strategy is to split what we are calculating in terms of majority and minority. The basic idea is to calculate the major portion first, sideline the minority, and compute it later.

2. Process Mapping: It is about deciding whether to use the push or pull approach. The same problem can be approached from the consumption and production side.
For Example, while finding the amount of chocolate used in India in a day.
We can have 2 approaches for the same.

- Consumption side: Estimating the number of end consumers (to avoid double-counting) and the units consumed by each consumer.
- Production side: Estimating the amount of cocoa produced, chocolate factories in a city and extrapolating it by appropriating the share based on the population of the country.

3. Data-centric Approach: Appropriating the right numbers in the structure when it is ready. To find numbers of an unknown category; you could assume a safe figure by considering a similar category.

4. Layout and then fill the numbers: Prepare a comprehensive exhaustive layout for the guesstimate and then start filling in the numbers.

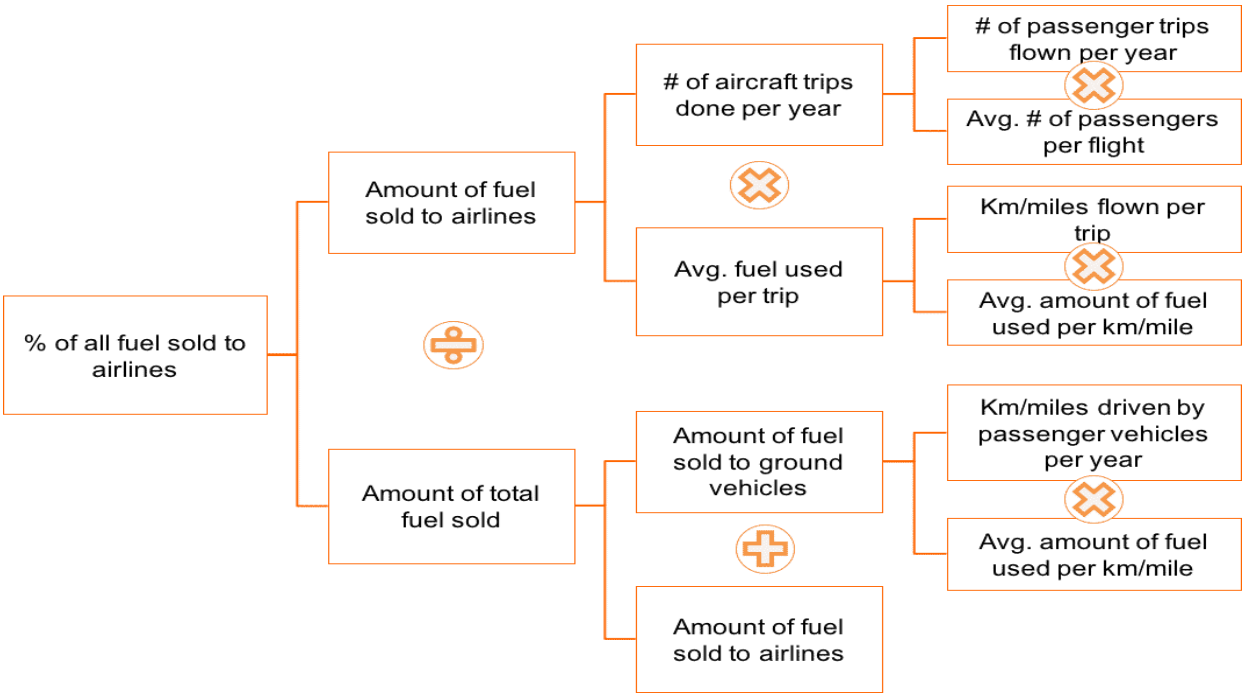
5. MECE (Mutually Exclusive and Collectively Exhaustive): Ensure that the buckets do not overlap with each other and are comprehensive when taken together.

Types of Guesstimates based on Approach to the Solution:

1. Household Approach: The category let's say, Cars are bought as a household purchase and hence we calculate the number of cars as per the number of households.

2. Population Approach: Example: A category like a pen is bought for individual consumption and is based on the number of people. Hence, we proceed with the guesstimate about the number of people.

3. Structural Approach: Example: To find the number of aeroplanes landing in India in a single day, the bottleneck would be the runway as it controls the entire operation.
This approach includes solving the guesstimate with a structure or logical breakdown, which is suitable for the question. Structure could be top-down, bottom-up, process mapping, layout centric, or critical comparison. The below mentioned figure describes a top-down structure approach, where you move from left to right, to estimate the market size.



Various filters to breakdown our problem in a systematic fashion:

1. Customer Segment:
- Rural-Urban (Geography-wise)
- Gender Split
- Age Split
- Income Split
- Willingness

For Example: How much sunflower oil is used in India in one year?
Here, a good idea would be to use a Geographic wise filter, as the consumption of sunflower oil is largely dependent on the region. Typically, we can break the region into 4 categories:
East
West

North
South
Each of these regions can be further broken down into Tier I, II, III & IV regions. Now, we can use this filter and proceed further by making logical assumptions backed by past experience.

Note that whatever, you have calculated for one region can be used for other regions by using a suitable multiplying factor.

2. Market Sizing:
- In terms of Volume, or
- In terms of Revenue

Whenever you are tasked with market sizing problems, ensure that you clear with the interviewer what exactly they are looking at, whether market size in terms of volumes or revenues. For example, the toothbrush market in India could be X units and Y USD. It is also helpful if you can clear the scope of calculation (Weekly, Monthly, Quarterly, Yearly, etc.)

3. Type of Product/Service
Usage of household products like TV, refrigerator, cars is completely different from individual products like toothbrushes, soap, cigarettes, etc. Whenever you are presented with a guesstimate, ensure that the first step after identifying the target customer is understanding the usage of the product. With regards to services ensure that you consider different scenarios, for example, to find the number of movie tickets sold in India every day, you need to consider the occupancy rate, which will be higher on weekends and for night shows.

4. Supply and Demand Side:
If it is required to find the number of petrol pumps in Bangalore, you need to consider the total number of vehicles in Bangalore and the number of filling pumps at a particular station. These two parameters together will be useful to estimate the total number of petrol pumps. Always remember that either demand constraint or supply constraint will drive calculation.

5. Types of product:
- Household products
- Individual products

Usage of household products like TV, refrigerator, cars is completely different from individual products like toothbrushes, soap, cigarettes, etc. Whenever you are presented with a guesstimate, ensure that the first step after identifying the target customer is understanding the usage of the product. With regards to services, ensure that you consider different scenarios, for example, to find the number of movie tickets sold in India every day, you need to consider the occupancy rate, which will be higher on weekends and for night shows.

6. Services:
- Peak and non-peak hours

- Weekends and Weekdays

Considering these factors provides additional depth to the calculations, which many of us tend to neglect. These factors are especially important for services (restaurants, movie theatres, malls) where demand is a function of day and time. To make the calculation more straightforward, you can consider two factors- peak and non-peak demand.

Elements of a Typical Solution

1. Clarify the question statement, have no ambiguity
2. Ask preliminary questions – gather information but do not try to solve the case
3. Reiterate objective function – be precise
   Note: In objective function, independent variable is what is asked to estimate in the question and dependent variables are the set of levers which are directly linked to independent variables. We try to solve Guesstimates by first designing a objective function and later work with interviewer to solve all the dependents variables
4. Take time to think
5. Give a detailed operating system with a possible hypothesis
6. Proceed with a framework (give mini summaries), ask questions to gather data
7. Understand what is the current set-up of the company
8. Provide conclusion and recommendations

Final Tips

1. Structured Thinking: The importance of this point can't be stressed enough – as this lays the edifice for solving a guesstimate. This is a trait, which is appreciated by the interviewer.
2. Sharp Communication: Remember the 3V's – Voice, Veracity, and Vocabulary. These traits come into play while you present your solution to the interviewer.
3. Smart Creativity: Build your own stories for making small assumptions, but be logical.
4. Know some basic facts/ data
5. Assumptions always take buy-in from the interviewer for your assumptions and clearly explain the rationale for the same.

Useful Data for Estimation

Population segmentation by income

All India (Avg household* income)

-Middle Class (> Rs. 16,000 pm): 30%
-Lower Middle Class (Rs. 8,000 - Rs. 16,000 pm): 40%
-Below poverty line** (< Rs. 8,000 pm): 30%

*Avg household size in India is between 4 and 5
**Poverty line defined as living below $1.25 per person per day
Note, numbers 8,000 and 16,000 translate to roughly 1 Lakh and 2 Lakhs respectively on an annual basis.

Avg household income of Mumbai/Delhi/Chennai/Other Metro cities

-Upper Middle Class (>32,000 pm): 10%
-Middle Class (16,000-32,000 pm): 30%
-Lower Middle Class (8000-16,000 pm): 40%
-Below poverty line (<8000p month): 20%

*Avg household size is 4 in Urban India

Consumption patterns in India
Avg. Consumption rate in India ~75% of income (varies with income segment)
Avg. Savings rate in India ~25% of income (varies with income segment)

Avg Discretionary spending rate is 25% of total consumption

Avg. Consumption in India (varies with income segment)

Segment	**Share of income**
Food	40%
Housing	10%
Transport	15%
Education	5%
Health	5%
Discretionary spending	25%

The above is meant as a simple guideline, good enough for case interviews and not an accurate representation. As income increases essential spending on Food decreases, while discretionary spending increases.
You do not need to remember the exact numbers, just remember expenditure on food is on average around 40%.

Population pattern in India

India: 1.25BN (can be taken as 1BN for ease in the calculation)
50% of the Indian population lies below the age of 25.

World		
Total Population	7,300,000,000	7.3 billion
Gender		
Male	50%	3.65 billion
Female	50%	3.65 billion
Continents		
Asia	60%	4.3 billion
Africa	15%	1.09 billion
Europe	10%	730 million
Latin America	10%	730 million
North America	5%	365 million
Age Groups		
0-5	10%	730 million
6-14	20%	1.46 billion
15-35	40%	2.9 billion
36-60	20%	1.46 billion
Above 60	10%	730 million
Income Group		
Poor	15%	1.09 billion
Low	55%	4 billion
Middle	15%	1.09 billion
Upper Middle	10%	730 million
High	5%	365 million

India		
Total Population	1,300,000,000	1.3 billion
Gender		
Male	50%	650 million
Female	50%	650 million
Rural Urban Divide		
Rural	70%	910 million
Urban	30%	390 million
Major Cities		
Delhi	1%	13 million
Mumbai	1%	13 million
Age Group		
0-14	30%	390 million
15-35	40%	520 million
35-60	25%	325 millon
Above 60	5%	65 million
Marital Status		
Married/Others	50%	650 million
Never Married	50%	650 million

Source: https://insideiim.com/how-to-solve-guesstimates-in-an-interview , Case Interviews Cracked , https://casereads.com/how-to-solve-guesstimate-questions-with-12-practice-questions/ , Decode and Conquer Lin, Lewis C

Solved Examples

1. Schools in Delhi

Preliminary Questions:
Should I consider both private and public schools?
- *Yes*

Can we eliminate dummy schools that are used by private coaching centers?
- *Yes*

Should I include play schools?
- *No, senior-secondary and primary schools only.*

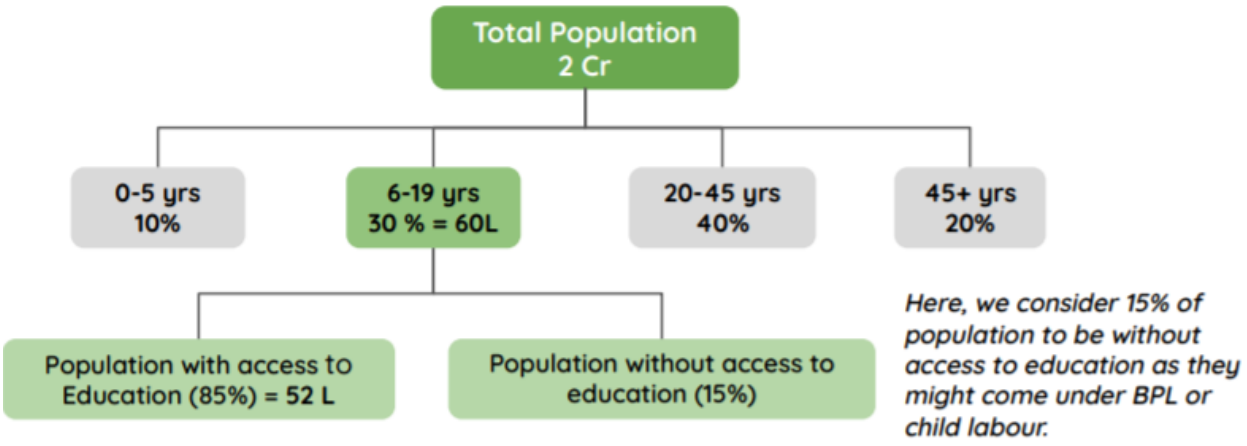
Assumptions
- School going age would be 6 to 19 years

- Population of Delhi ~ 2 Cr
- There are 10% large sized schools (Private) 40% medium sized schools (Govt + Private), and 50% small sized schools (Govt + MCD)

Overall Strategy
(1) We can estimate the number of students in a school: Number of Grades x Avg Number of sections x Avg Students per class
(2) Calculate the total number of students in Delhi by age segregation

1) Large Sized School: 14 Grades (Nursery-12th) x 10 x 40 ~ 5.5k Students
2) Medium Sized School: 12 Grades (1st-12th) x 3 x 30 ~ 1k Students
3) Small Sized School: 10 Grades (1st-10th) x 2 x 20 ~ 400 Students Therefore, Average number of students in a school in Delhi: (0.1 x 5,500 + 0.4 x 1,000 + 0.5 x 400) = 550 + 400 + 200 = 1150

Now, Delhi's population can be segregated on the basis of age demographics



Therefore - # schools in Delhi = Total number of students in Delhi / Average number of students in a school in Delhi = 52,00,000 / 1,150 ~ **4500 Schools**
Home Task -

1) **Cars Crossing Delhi-Gurugram Toll Everyday**
2) **Crocin Tablets sold in Delhi in one day**
3) **Emojis sent in India each day**

One to One Interaction -

2. Burgers Sold at McDonald's

Could you estimate for me, the number of burgers a McDonald's outlet sells in a day?

Preliminary questions

Can we assume this to be an average McDonald's as opposed to one in a specific location like an Airport?
- *Yes.*

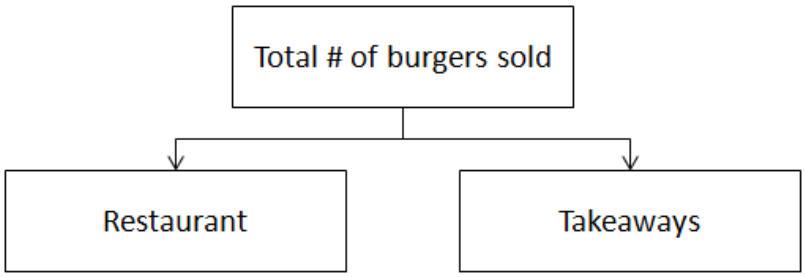
Do we include sales from takeaways as well?
- *Good question. It might be useful to include takeaways.*

Do we assume that the outlet serves other items like wraps/puffs etc.? We will discount the people having wraps/puffs instead of burgers.
- *Assume that burgers are the only item on the menu apart from fries and coke.*

It is very useful to scope out the problem correctly. This has 2 advantages.
(1) You need to identify the constraints you are working under so as to analyze only what is required by the client and move through the case interview faster.
(2) The interviewer notes that you have thought about various aspects of a case. Although he/she might not ask you to delve into the nitty-gritties, he/she is assured that if you're hired and assigned to a case, you will cover issues exhaustively.

<u>Overall strategy</u>

Here is what I want to do.
To simplify this problem, I would like to estimate the sales in (1) the restaurant and (2) takeaways, separately.



(1) Restaurant

<u>Strategy</u>
I would like to approach this problem from a supply point of view. I want to add that supply in this context means not the burgers that can be manufactured but the maximum customers that can be seated in McDonalds on a given day. I will further try to understand their consumption patterns to arrive at total burgers sold.
I'm assuming that an average McDonalds has about 50 seats and is open from 10 am to 10pm. The consumption patterns of burgers are different throughout the day

The total number of burgers = (# hours) x (# people per hour) x (# burgers per person)

We have used mathematics here to ensure MECE segmentation

.
The total number of burgers = (# hours) x (# of seats occupied) x (# people per seat per hour) x (# burgers per person)
The total number of burgers = (# hours) x (Total # of seats) x (average % occupancy) x (# people per seat per hour) x (# burgers per person)

Let's say every person eats at McDonalds for 20 minutes. There are 3 (60min/20min) people occupying a seat every hour. Also, in my experience since burgers in India are slightly smaller, I will assume 20% of the people eat 2 burgers and 80% of eat 1 burger. That is an average of

0.2*2 + 0.8*1 = 1.2 burgers per person per sitting.
The total number of burgers = (# hours) x (50) x (average % occupancy) x (3) x (1.2)
= (# hours) x (% occupancy) x (180 or approx. 200)

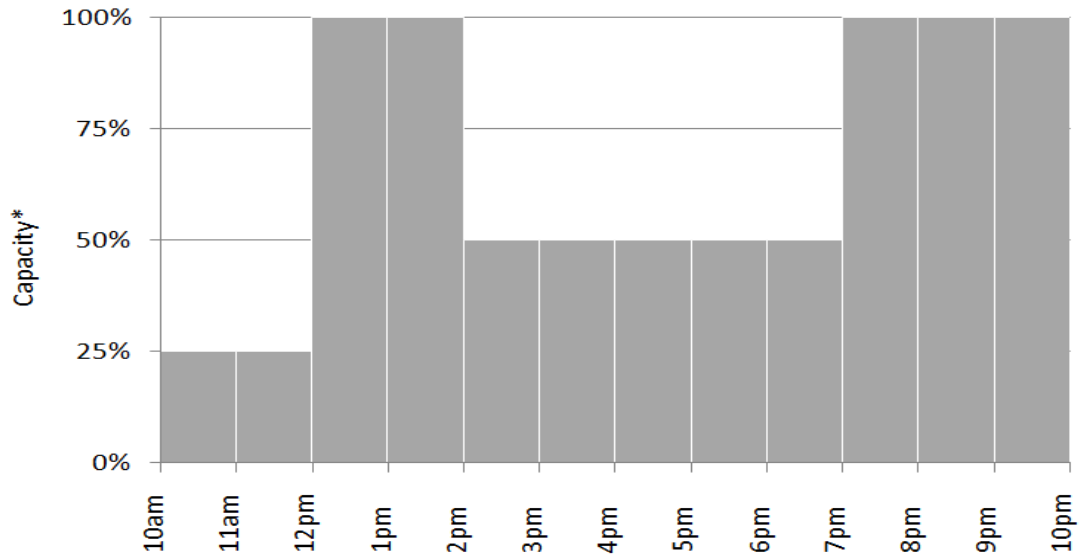
Since the occupancy varies according to time of day, I would like to do the math separately for each hour. Do you think that works?
- *Sure. That seems reasonable.*

For simplicity's sake, let's take 3 scenarios: 100% occupancy (high traffic), 50% occupancy (medium traffic), and 25% occupancy (low traffic).

High traffic: Lunch and dinner hours
Medium traffic: Post lunch and early evening hours
Low traffic: Morning hours

The kind of insight demonstrated by the candidate above is what you want to aim for in guesstimates. Not only should it be important but something that can easily be incorporated in your solution.

Here is a graph representing the traffic of people at McDonald's by the hour.



**100% capacity indicates a completely filled McDonalds*

Calculation: (Following from the last mathematical equation)

High traffic: 5 hours * 100% occupancy * 200 burgers = 1000 burgers
Medium traffic: 5 hours * 50% occupancy * 200 burgers = 500 burgers
Low traffic: 2 hours * 25% occupancy * 200 burgers = 100 burgers

Total number of burgers sold in a day in a restaurant is 1600 (or ~1500 burgers).

(2) Takeaways

The takeaway counter has a queue during high traffic hours and it will be useful to bring in my own experience at these counters to estimate the time that each exchange takes.

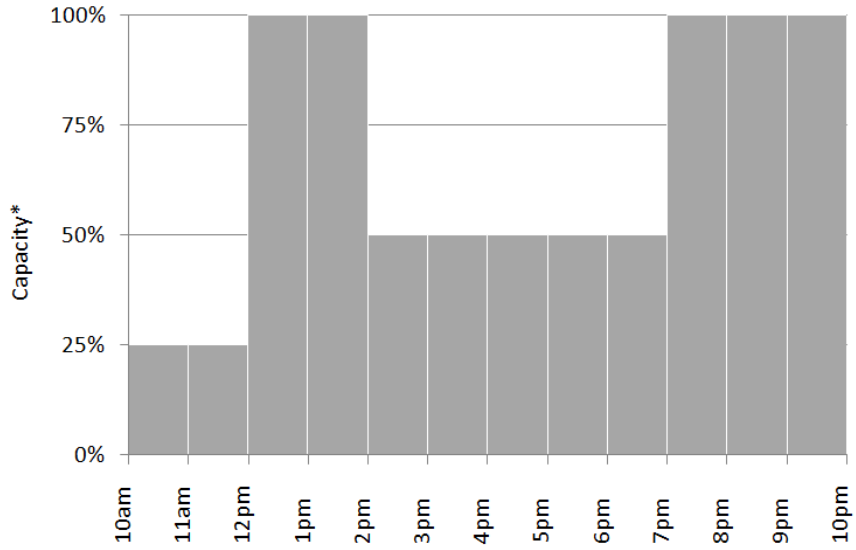
In my view,
Time taken for every person = 75 s = 1.25 min

# of burgers in an hour = (# of people in an hour) * (# burgers per person)

I have assumed that the takeaway orders are slightly more than the restaurant orders.

# of burgers in an hour = (60/1.25) * (2) = ~100 burgers

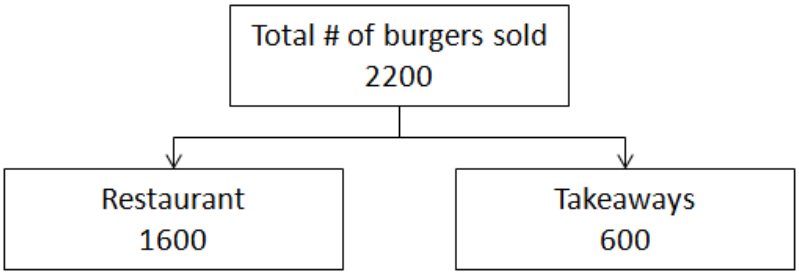
Assuming that the traffic of people is the similar for takeaways as well,



**100% capacity indicates a completely filled McDonalds*

High traffic: 5 hours * 100% traffic * 70 burgers = 350 burgers
Medium traffic: 5 hours * 50% traffic * 70 burgers = 175 burgers
Low traffic: 2 hours * 25% traffic * 70 burgers = 35 burgers

Total number of burgers sold in a day in a restaurant is 560 (or ~600 burgers).



Hence, the total number of burgers sold every day at an average McDonald's is about 2200 burgers
- Impressive. Is there anything else you would like to add?

Yes. I would like to do a quick sanity check.

- *Okay. How might you do that?*

I think the bottleneck during high traffic hours is the supply i.e. the manufacturing rate of the burgers themselves. The kitchen would be running at nearly full capacity during these hours.

Supply = Demand during peak hours.

<u>Demand (as per our calculation)</u>

Demand of burgers during high traffic time = 270 (restaurant: 200; takeaways: 70) burgers per hour = ~4-5 burgers per minute.

Supply

Keeping in mind that the McDonald's model is a made-to-assemble* one,

Time required to make a burger = (1) Sourcing components + (2) Heating + (3) Assembly + (4) Delivery

***A made-to-assemble model is one where the individual components are ready or 'made' and require only assembling to make the finished product**

Time required to make a burger = 5s + *5s + 10s + 10s = ~30s

(*Assuming 2 crate of patties (each having 30) take about 5 minutes to cook.)

Hence, every kitchen employee makes 2 burgers per minute.

Assuming 3 people working in the kitchen during high traffic hours, the outlet produces ~6 burgers per minute or which is in the same ballpark as the demand. I concede that due to incomplete knowledge about the industry, I may have made some errors in assumption.

- That is just fine. I like your approach.

3. Market Size of ACs in Mumbai.

Preliminary Questions

There are two types of air-conditioners- window and split. Shall I include both?

- Yes.

And should I include objects like air-coolers?

- No, you need not.

In what unit would you want me to estimate the market size- #ACs, INR, etc.?

- Good question. Estimate the market size in terms of tonnage. A 1 ton AC can typically cool a small room.

Should I include second hand ACs as well?

- Interesting, but no. Include only first-hand ACs in the market size.

Air Conditioners are used for residential cooling and commercial cooling in Mumbai. Residential cooling involves cooling in home and apartments. Commercial cooling would involve Office Spaces, Malls, Restaurants, Hospitals, Schools, etc. Shall I calculate the market size including all these?

- For simplicity, just calculate Residential and Office Space cooling. We can neglect the rest.

Overall Strategy

We need to calculate the market size for air-conditioners in tons. The market size would involve the #ACs required to replace old ones along with the additional demand for new ACs.

Market size in tons = Demand to replace old ACs + Demand due to growth in market size
= (#ACs in tons) / (Avg. life of an AC) + (#ACs in tons) x (Growth rate)

Avg. Replacement Demand = Current # Products/ Avg. Life cycle of product.
If the life cycle of a product is Y years on average, then in Y years' time each of these products would be replaced by newer ones. Then we can say, Average replacement demand for any particular year is (Current # Products/Y).

We can take the average life of an AC to be 10 years. Can we take the growth rate in absence of any other information as the growth in GDP of India which is about 5%?
- *Take the growth rate to be 10%.*

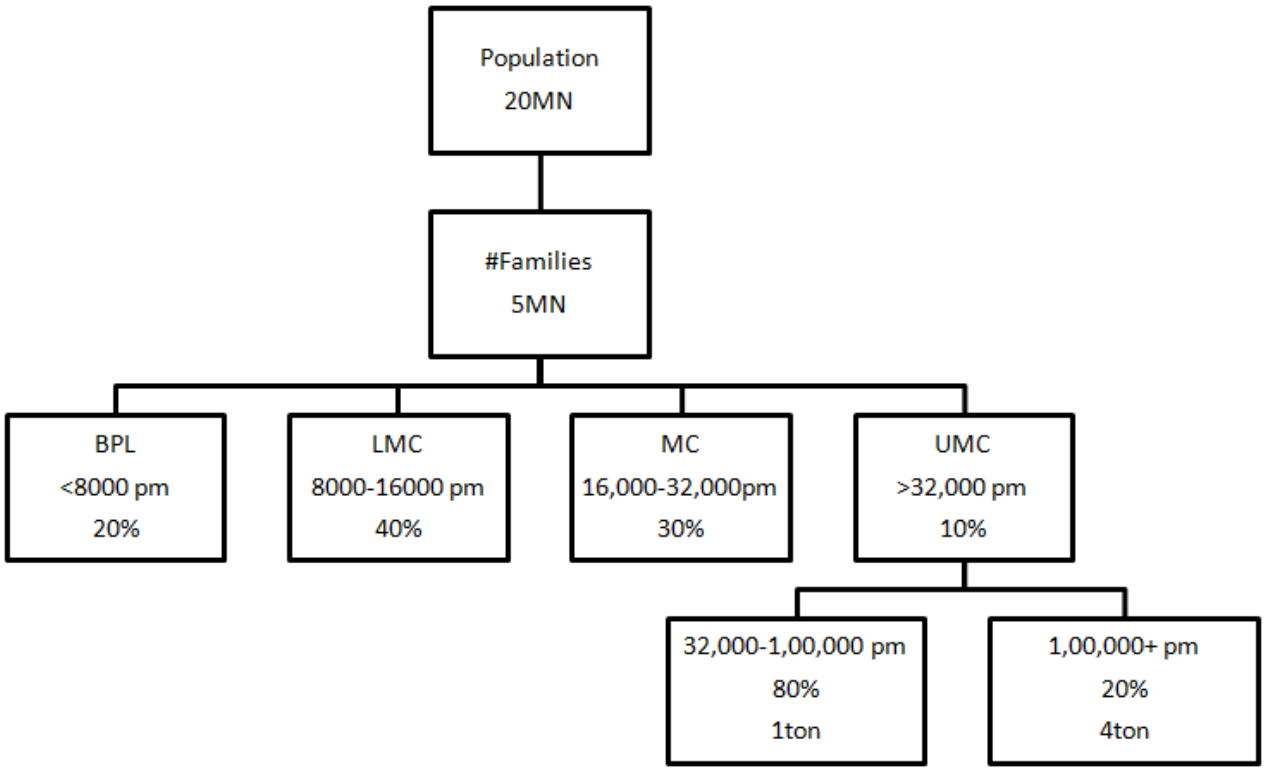
Alright, what remains to be calculated is the (#ACs in tons) presently in Mumbai. We are looking at demand in
● Residential Segment
● Office Space Segment

For the Residential segment:
#ACs in tons = (#Families who can afford ACs) x (Avg tons of ACs per family)

For the Office Space Segment:
Offices typically have centralized ACs and the tonnage of the Central AC decides how much area can be cooled.
#ACs in tons = (Total office area) x (Tons required per unit area)
= (#People working in AC offices) x (Office area per person) x (Tons required/ unit area)
- *Sounds good. Go ahead.*

Residential segment
We can take the average family size to be 4 in Mumbai. Though there are families of bigger and smaller sizes, we will use this simplification. Hence for a population of 20MN we get 5MN families.

The next aspect we need to look at is the affordability of an Air Conditioner. An AC costs roughly INR 25,000 to 40,000 depending on the tons.
A Middle Class (MC) family has an income in the range of 16,000 to 32,000 pm. The average household income would around 20,000pm. **Note that the average is skewed towards the lower limit.** This would probably be the salary of junior most clerks in offices. Given the AC cost is more than a month's salary, I do not think that the Middle Class would be able to afford ACs. Obviously the Below Poverty Line (BPL) and Lower Middle Class (LMC) families would not be able to afford it as well.



Now we need to see the Upper Middle Class (UMC) families. Their income is greater than 32,000 pm. The average income of this segment would be roughly 50,000 pm or about 6 lakhs per annum. The average UMC family would live in a 2 BHK home. I think such a family would be able to afford at least a 1 ton AC.

In this segment there would be families which can even afford up to 3ACs, two 1 ton ACs for each bedroom and a 2 ton AC for the living room. This is the case for my family, and our average household income would be around 100,000pm. Based on this we can split the UMC into two segments.

- Income between 32,000pm-100,000pm. Let's assume this will be 80% of the UMC segment. We will assume on average a family in this segment owns a 1 ton AC.
- Income greater than 100,000pm. Based on our assumption this will be 20% of the UMC segment. We will assume the average family in this segment owns 4 tons of ACs

Hence we get,

#Tons in residential segment= 5MN x (10%) x (80%) x (1 ton) + 5MN x (10%) x (20%) x (4 tons)

=400,000 + 400,000 =800,000 tons.

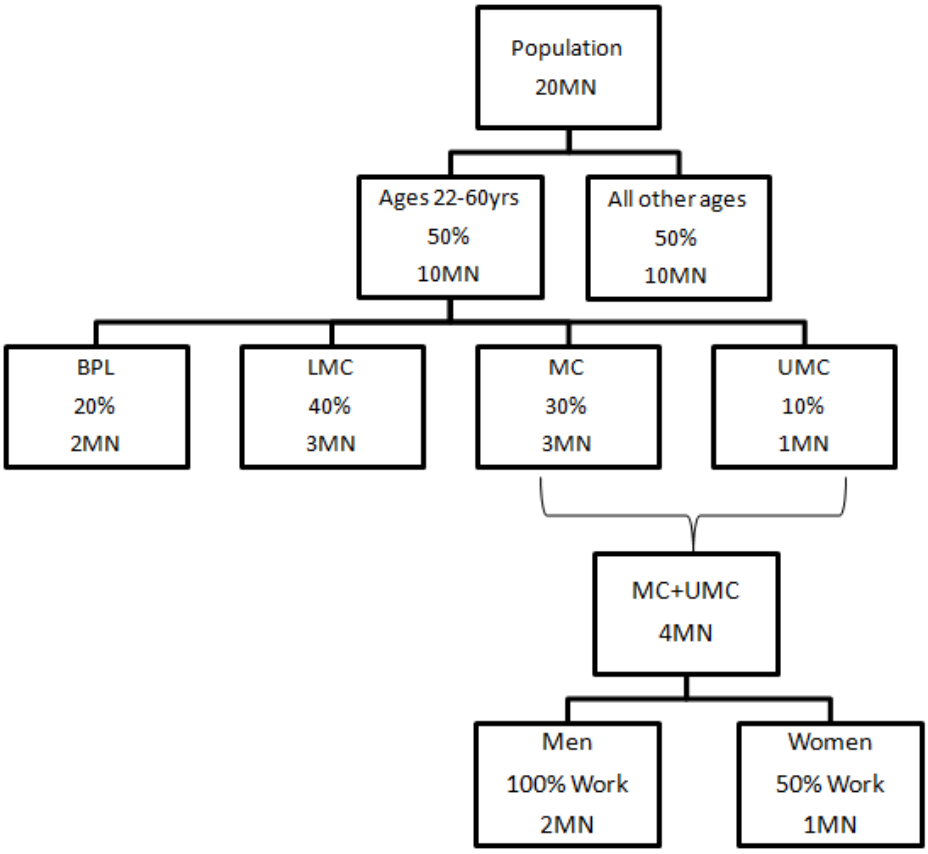
<u>Office Space Segment</u>

For the Office Space Segment

#ACs in tons= (#People working in AC offices)x(Office area per person) x(Tons required per unit area)

The people working in air-conditioned offices typically would have white collar jobs. Though there will be cases where white collar professionals do not work under ACs and blue-collar workers have air conditioning, we will neglect these exceptions. (White collar employees are those which typically work

in offices, they are accountants, bankers, bureaucrats, etc. Blue collar employees would be mechanics, electricians, manual laborers, etc.)



The people with white collar jobs would be part of the Middle Class and Upper Middle Class in the age group of 22-60. Assuming an equal spread of the population from 1-80 yrs., roughly 50% of the population will lie in the 22-60 bracket.

We can assume that all men and half the women would be working.
This gives us a total of 2MN men + 1MN women = 3MN people.

In a typical office, people work in cubicles. Apart from the area of a cubicle, an air-conditioned office has several common areas such as receptions, washrooms, lunch areas. If we were to divide the total area of an office with the total number of employees, we can say that an office area roughly the same as a ‘bedroom’ would be equivalent to two employees. A ‘bedroom’ typically has an area of 200 sq. ft. Hence each employee is equivalent to 100 sq. ft. area in an office. Also we will require a 1 ton AC to cool a room of 200 sq. ft.
Hence,
#ACs in tons= (#People working in AC offices) x(Office area per person)x(Tons required per unit area)
= (3MN) x (100 sq. ft.) x (1 ton / 200 sq. ft.) = 1.5MN tons

Combining the office segment with residential segment, we get the total AC tons to be 2.3MN.

Market Size in tons = (#ACs in tons) / (Avg. life of an AC) + (#ACs in tons) x (Growth rate)
= (2.3MN tons) / (10 yrs.) + (2.3MN tons) x (10% per year)
= 460,000 tons / year
- *Great. We can close the case here.*

4. Number Refrigerators Sold in India every year.

Preliminary questions:

Should we consider domestic refrigerators only or should we include commercial refrigerators too?
- *For this question, include only domestic refrigerators.*

I would like to approach this question from the demand side.

Total population of India = 130 crore
Let us take a number of members in a family = 4 ; Total Households = 130/4 = 32.5 crore

For Metro/Tier I class households
Percentage of poor: middle class: rich = 20%:60%:20%

For Tier II/III class households
Percentage of poor: middle class: rich = 40%:50%:10%

For Village Households
Percentage of poor: middle class: rich = 60%:38%:2%

Market Penetration

	Metro/ Tier - I	Tier- II/III	Villages
Poor%	0	0	0
Middle class %	50%	40%	0
Rich %	100%	100%	100%

Absolute Refrigerators Numbers = Sum of Penetration * No of households in each segment

	Metro/ Tier - I	Tier- II/III	Villages
Households	0.2*32.5 = 6.5 Cr	0.15*32.5 = 4.875 (Let's take it 5)	0.65*32.5 = 21.125 (Let's take it 21)
Refrigerators Penetration	0.5*0.6+0.2 = 50%	0.4*0.5+0.2 = 30%	2%

Total	0.5*6.5= 3.25	0.3*5 = 1.5	0.02*21 = 0.42

Total number of refrigerators = 3.25+1.5+0.42 = 5.17 crore

As the question asks the number of refrigerators sold every year, we have to take care of following points too:
i) People buying refrigerators due to replacement (as their old machine is not working)
ii) People buying refrigerators as their first refrigerator

Let's take the average life of Refrigerators= 10 years
Assumed growth rate = 5%

Refrigerators sold in a year = refrigerators due to replacement + New Demand due to growth
= (Total refrigerators (i.e. market size) / Avg Life) + (Growth rate * Total machines (i.e. market size))
= (5.17/10)+(5.17*0.05) crore = 0.775 crores = 77.5 lakhs

5. The fleet size of Air India

Preliminary Questions

Do I need to calculate both international and domestic fleet size?
- *Count only the domestic one as of now.*

Should I consider only active aircrafts or also ones under maintenance
- *Only the active ones.*

Overall Strategy

I would like to first estimate how many flight routes Air India operates in. Then we can see what would be frequency of each route and accordingly find the no. of aircrafts required to service the route. Finally we can sum up the required aircrafts for each route and get the result.
- *Sounds good, go ahead.*

There are two types of cities
1. 6 Metros: Mumbai, Delhi, Kolkata, Chennai, Hyderabad, Bangalore
2. 30 Tier-two cities

This would lead to three types of direct routes
1. Metro to Metro: 15 routes

2. Metro to Tier two: Total of 6x5=30 routes.

We can assume that every metro has 5 tier-two orbital cities, for example there is no direct flight from Mumbai-Chandigarh. It is routed via Delhi. Hence Chandigarh is an orbital city of Delhi. We will assume there are direct flights only between a metro and its orbital tier-two cities, not between a metro and different metro's orbital tier-two city.

3. Tier-two to Tier-two (direct flights are rare, since they usually routed via Metros and hence accommodated in previous two types of routes)

There may be some flights which have multiple stopovers. We can consider them to be equivalent to two flights (if there is one stop over) under the types of routes mentioned above and continue with the analysis.

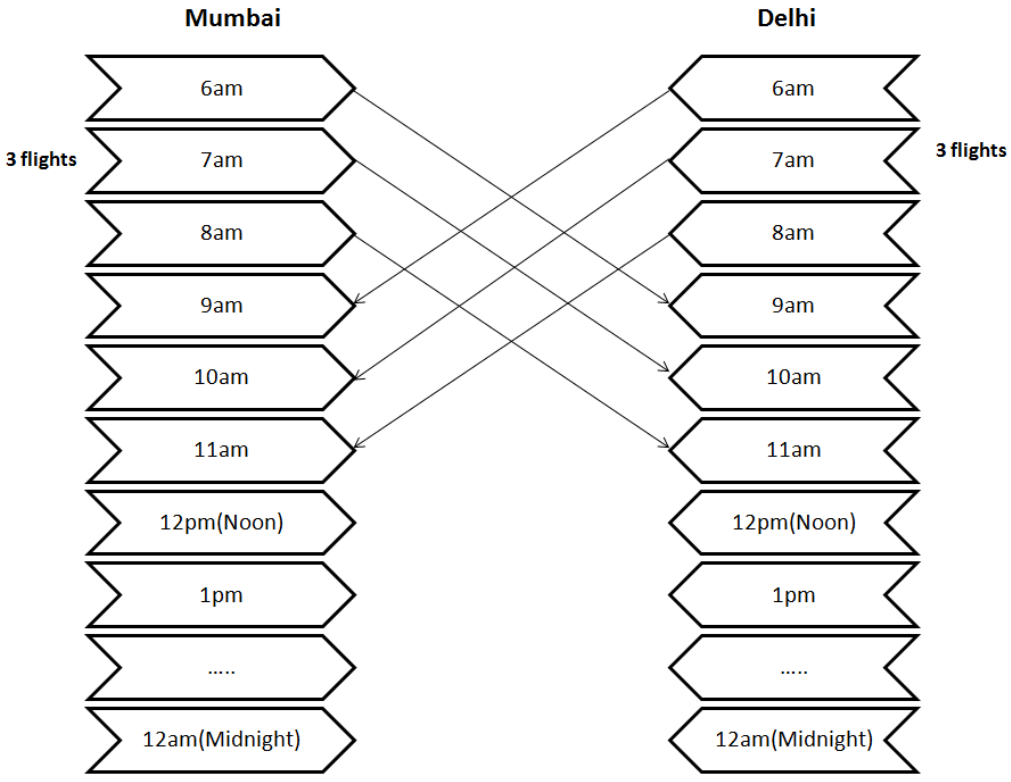
<u>1. Estimating # flights it would take to cover a Metro-Metro route</u>

Let's assume flights can run between 6am in the morning to 12am in the night.

If we take a high intensity metro route like Mumbai to Delhi, based on my experience, there is a flight of Air India departing every hour.

If I assume that a flight leaves from Mumbai at 6am. It reaches Delhi at 8am. Then there is a 1 hour turnaround time which includes maintenance of the aircraft. Then the flight leaves from Delhi at 9am. Similarly after reaching Mumbai, the flight is ready to depart at 12pm (Noon).

At 6 am simultaneously a flight departs from Delhi, reaches Mumbai at 8am and after a 1 hr. turnaround time departs from Mumbai at 9am.



Totally 6 flights required to cover Mumbai-Delhi route to achieve frequency of 1 flight departing per hour

This would mean that we need 3 flights departing from Mumbai at 6am, 7am & 8am and 3 flights departing from Delhi at 6am, 7am & 8am to achieve a 1 hour flight frequency from 6am to 12pm (Midnight). Therefore a route like Mumbai-Delhi requires 6 flights.
However all Metro to Metro routes are not this busy. Mumbai-Delhi is the busiest route. If we take Kolkata to Hyderabad, it is a less busy route and the flight frequency would be more like 1 flight per 3 hours by Air India. This would mean a requirement of only 2 flights to cover the route.

On an average we can assume a requirement of 4 flights per metro route leading to an average frequency of -a flight every 1.5 hours.
This would mean 4X15=60 aircrafts on the Metro to Metro route.

2. Estimating # flights it would take to cover a Metro-Tier-two route
Coming to the 30 Metro -Tier-two routes. A typical route would be Delhi-Chandigarh.
These routes have low frequency. The journey time is less than an hour.

A flight departing at 6am from Delhi will reach Chandigarh at 7am, would depart from Chandigarh at 8am, would reach Delhi at 9am and would again leave from Delhi at 10am.

Hence with one flight we can achieve a flight frequency of 4 flights per day. Based on my experience **(You can ask the interviewer for his experience, makes the interview conversational)** even 4 flights might be too much and Air India may be having just 3 per day to save on operating expenses. Additionally it's possible that one air-craft might be catering to two low frequency Metro--Tier 2 routes, however we will neglect that and assume each route will have one dedicated aircraft at least. That leads to 30x1=30 Aircrafts.

Hence we have a total of 60+30=90 aircrafts of Air India meant for domestic routes in active service.
- Alright, good job.

6. No. of taxis in Mumbai

Preliminary Questions

What kind of taxis, call cabs like Meru and OLA? Or the spot taxis like Black and Yellows?
- Calculate the spot taxis, only, mainly the Black and Yellows.

There is an overlap of demand between call cabs and spot taxis. However, let’s assume it is insignificant since call cabs are used for longer distances like going to airport etc., and have a more premium income segment.
- Fair enough.

Overall Strategy

We can use a Demand based approach.

#Spot Taxis = (Total #Taxi fares per hour) / (Avg. fares per taxi in a hour)

The above analysis can be done for any hour in the day. However we should choose the hour on the basis of ease of computation of (Total #Taxi fares) and (Avg. fares per taxi) in that hour.

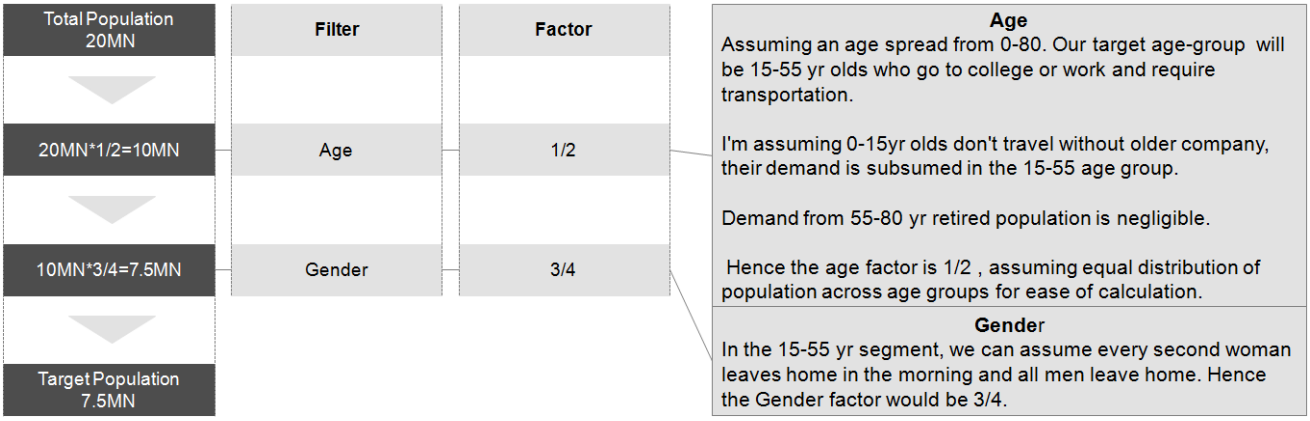
For this let us choose the morning time, since we know most people need to leave home to go to work or college at that time. This is better than choosing some other time like afternoon/evening where demand can come for a wide variety of purposes (like going for shopping, theatre) and will make it difficult to compute. We can compute the

(Total # Taxi fares) in the morning time (7am-12pm) and divide by number of morning hours to arrive at (Total #Taxi fares per hour) in the morning.

We need to consider the following factors when calculating (Total # Taxi fares) in the morning

- Population
- Income Split
- Age split
- Gender
- Alternatives: Buses, personal cars, bikes, walking, auto-rickshaws, other vehicles
- Region: Auto-Rickshaws are not allowed in South Bombay, so use of taxis is greater there
- No. of passengers in a taxi at a time
- Journey time

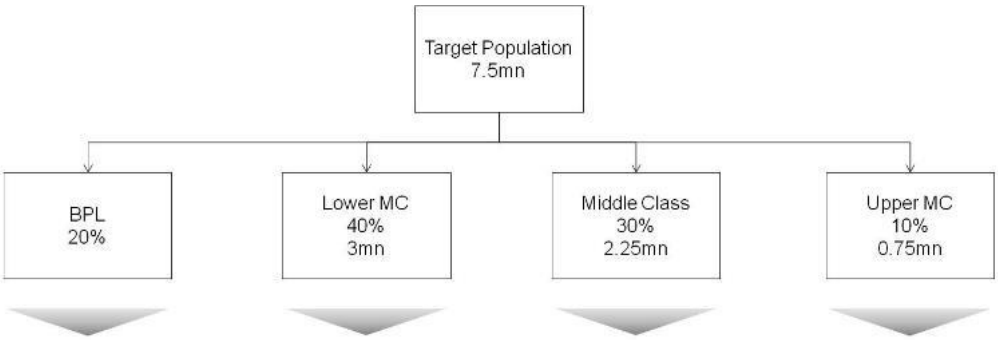
Let us begin the estimation



Of this target population of 7.5mn, we can now find how many require transportation through autos or taxis. Later we can compute, how many of them prefer taxis.

This step may be counter-intuitive. Try to directly go ahead with finding taxi demand instead of (taxi + auto) demand. You would find (as we did) the need to do the income-split analysis, below, separately for South Bombay and Rest of Mumbai. The approach given does the income-split analysis together, and then applies a factor for market share of taxis within (taxis + autos). The

lesson here is that in the middle of a case you can pull back, tweak your approach for simplification purpose, and then resume forward.



Characteristics	Cannot afford taxis/autos	Use buses for longer distances, walk shorter distances. They car pool up to 3-4 persons from local taxi/auto spots to workplace hubs.	Do not own a personal car but can better afford taxis/autos. Less dependent on car pooling, avg of 2 persons.	Have a private car so don't always need a taxi/autos, but also don't use buses too Can afford to not car-pool
% using taxis and autos	0	40%	60%	20%

This gives us a total of 1.2mn+1.35mn+0.15mn=2.7mn people requiring (taxis + autos).

We need to now apply a factor, for people only using taxis out of these.

In South Bombay, autos are legally not allowed, so people there will only use taxis. South Bombay roughly represents 20% of Mumbai's population. In the rest of Mumbai, most people use auto-rickshaws since their more easily available and cheaper. Let's say 20% of people outside South Bombay use taxis in the morning.

Hence % of people using taxis in the morning= 0.2*100%+0.8*20% = 0.36
Hence the #people requiring taxis in the morning = 2.7mn * 0.36 =~2.5mn * 0.4 = ~1mn

Now, we need to apply a factor for carpooling. This differs by income-segment. In the middle-class segment on an average two people would use a taxi at a time. Since this is the largest segment of demand, we will consider the average carpooling number to be 2.

Hence we get a demand for 1mn/2=500,000 taxis during the morning.

Now, let's say the demand is scattered over morning 7am-12noon, with peak-hour being 9am-10am constituting 30% of the demand.

During peak hour of 9am-10am, there is requirement of = (500,000)x(30%) =150,000 taxis per hour

An average fare time during the peak hour would be =Time with passenger + Time without passenger
=40mins+5mins = 45mins
Avg. fares per taxi in a hour = 1 hour / (Average fare time) = 60 mins / 45mins = 1.33
Hence the total # taxis= (Total #Taxi fares per hour) / (Avg. fares per taxi in a hour)
=150,000 / 1.33
~110,000 taxis

Sanity-Check
As a sanity check let's see what the ratio of taxis and passenger cars is:
We know population of Mumbai is 20MN. This will roughly equal to 5MN households.
We can assume that according to the income segment, the top 20% can afford cars.
This will give us about 1MN passenger cars.

The ratio of taxis to passenger cars would then be ~1:10 which seems to be in the right ball-park based on experience.

Note the above computation of passenger cars is not how you would do it if you were asked an entire guesstimate on it. However it is a useful 'quick and dirty' method for the sanity check, here.

7. Cost of painting the pillars of the metro lines in Delhi

The crucial step here is to estimate the number of pillars.
We try to visualize how any metro line looks like. It starts from a station and ends at another. There are turns, splits and intermediary stations.

The line itself may be underground, on-ground or above the ground

1. Underground and On-ground - Since the underground/ground level stations will not have any pillars, we will assume the number of pillars for them to be zero. The connections to the preceding and subsequent stations, however, will be taken into account

2. Pillars at station - will have more number of pillars, owing to the extra strength required for the stability of the station premises, therefore, we will assume 8 pillars for each station above ground level.

3. Turn/ split - The number of pillars increases for every turn/split(like the one after Yamuna Bank) in the tracks. We will assume the increase on the basis of the length of the track.

4. For every 100m, there is a pillar. Therefore, there are 10 pillars in each km.

Basis the above assumptions, the length of the various Metro Lines, and subsequently, distance between two stations is assumed to be as follows:

Line	**Length (L)** (in km)	**Number of stations (S)**	**Number of Underground /Ground Level stations (U)**	**Number of stations above ground level (A)**	**Distance between two stations (L/S)** (in km)	**Distance to be considered (A X L/S)**
Blue	52	44	4	40	1.2	48
Red	25	21	4	17	1.2	20.4
Yellow	45	34	25	9	1.3	11.7
Orange	22	7	5	4	3.1	12.4
Violet	23	18	6	12	1.3	15.6
Green	20	17	0	17	1.2	20.4

Now calculating th eno of pillars according to the DMRC data and our assumptions:

Line	**Number of pillars for each station above ground level (8 X A)**	**Increase in pillars due to turns/splits**	**Number of pillars for the distance to be considered (10 X A X L/S)**	**Total number of pillars**
Blue	320	20	480	880
Red	136	12	204	352
Yellow	72	0	117	189
Orange	32	0	124	156
Violet	96	8	156	260
Green	136	8	204	348
			GRAND TOTAL	2185

Now, once the number of pillars are calculated, we can estimate the cost of painting the pillars as well as the revenue. The following figures have been taken arbitrarily

- Assume the cost of painting each pillar to be Rs 1200.
- Assume the revenue from each advertisement placed to be 2500.

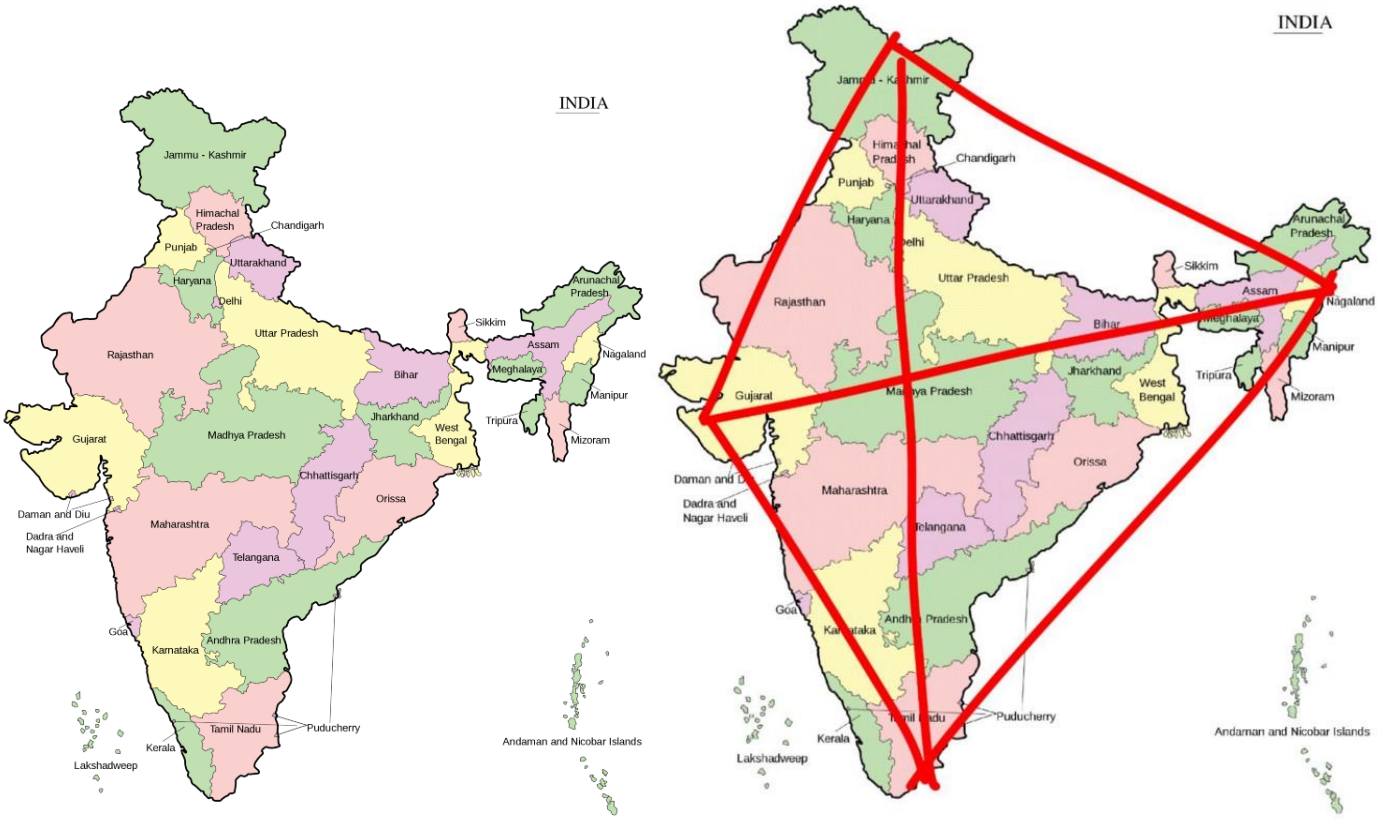
Number of pillars	Total cost of painting (1500 per pillar)	Revenue earned (2500 per pillar)	Total expenditure
2185	Rs. 32,77,500	Rs. 54,62,500	(Rs. 21,85,000)
		NET EXPENDITURE (FINAL ANSWER)	(Rs. 21,85,000)

Further specifications could have made the estimation more precise, following could be considered:

1. Assigning a higher revenue to popular areas (Cost of advertising would be higher) and lesser revenue for those less popular.
2. The above step would lead to the segregation of the metro line's distance into densely populated/less populated areas. Alternatively, we may also choose to list the most well known stations where the advertising cost is bound to be higher.
3. Inculcate the metro lines still in construction and their pillars.
4. We could have also assumed specific dimensions for the diameter and the height of the pillars, subsequently arriving at its area and multiplying the same with the cost of painting per square unit area.

8. The area of India

The first thing to come to your mind should be the map of India. Look at it carefully:



Looks like a rhombus
Area of a rhombus =(product of diagonals/2)

Exclusion: Area of the islands is considered to be insignificant

Since we need the diagonals for the area, we need to calculate the distance between Northernmost & the Southernmost points and Easternmost and Westernmost points.

(The best case scenario is that you know these values. If you don't you have a way out.)

Distance = speed x time

We simply need to find out something that involves speed and time — a mode of transport. Pick from — road, rail or air. Go ahead with trains.

The average speed of trains in India can be taken as 50 km/h.

(They move at about 60–80 km/h and then consider all the unnecessary stoppages. Also, ease of calculation with 50.)

Time taken to travel by trains from:

North to South: Jammu to Delhi = 10 hours
Delhi to Kanyakumari = 50 hours
Total = 60 hours
Distance = 50 X 60 = 3000 km

West to East: Ahmedabad to Delhi = 10 hours
Delhi to Itanagar = 40 hours
Total = 50 hours
Distance = 50 X 50 = 2500 km

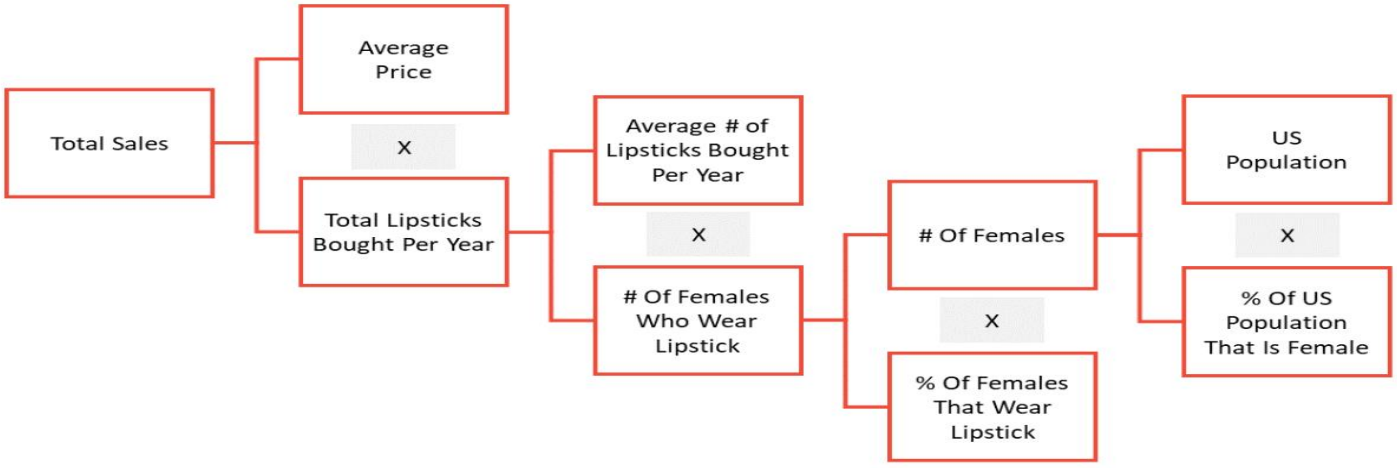
Therefore, area of India = (3000 x 2500)/2 = (7,500,000/2)

The guesstimated area of India is 3.750 million km sq.

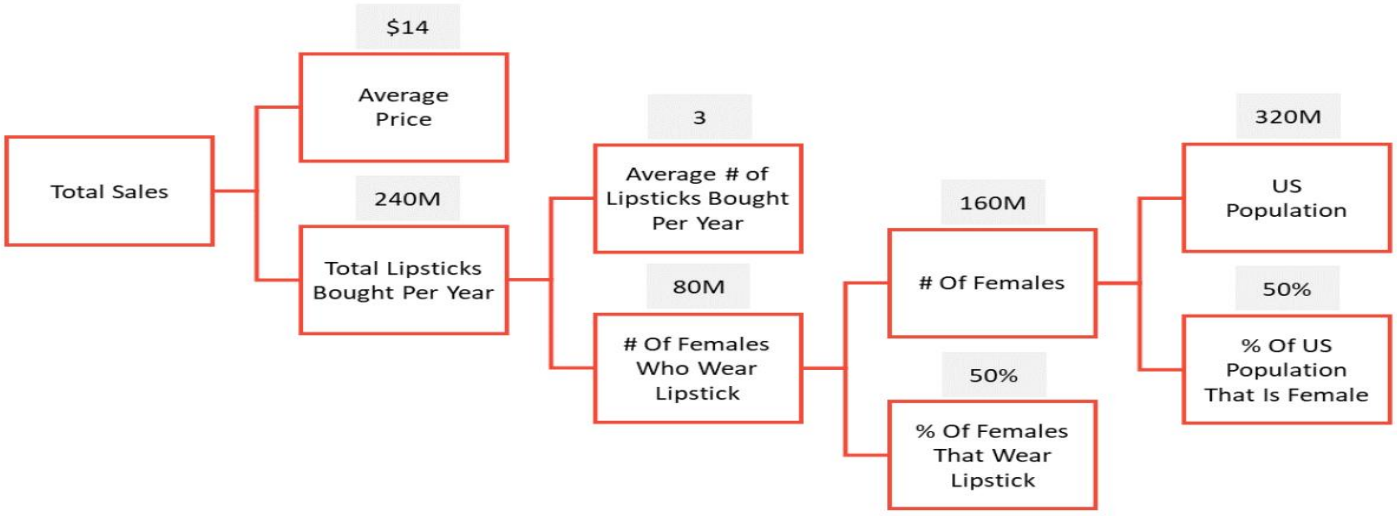
Sanity check: The actual area of India = 3.287 million km sq.
And this 14% error is completely acceptable.

9. Market size of lipstick in the US

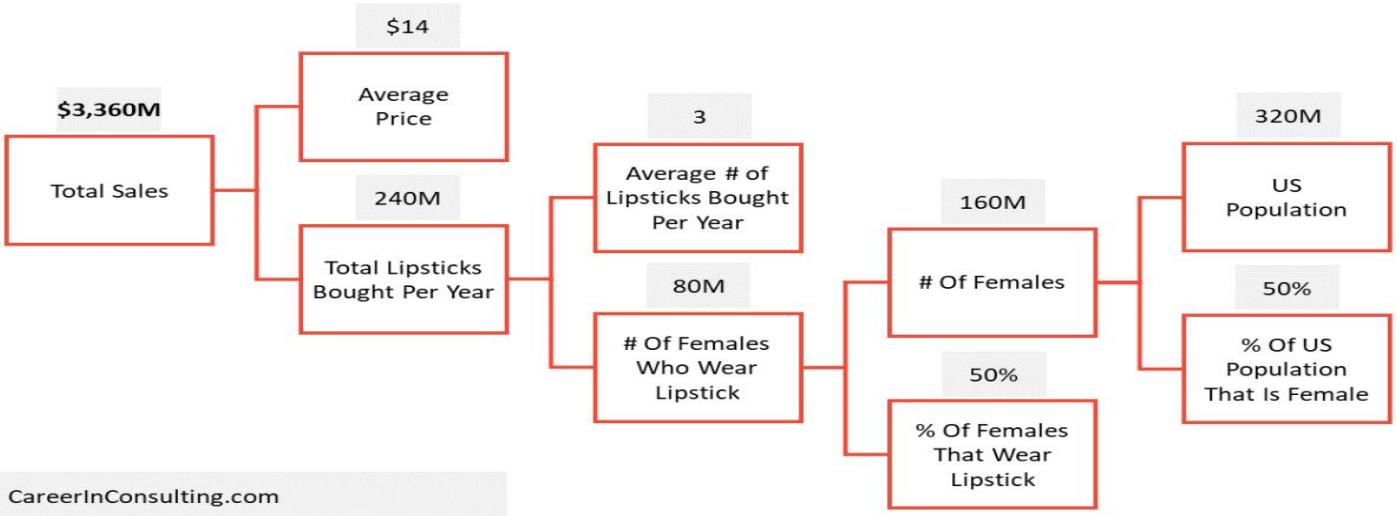
1. Breakdown The Problem Into Smaller Pieces



2. Make An Assumption For Each Of These Smaller Pieces



3. Recombine All The Pieces Together To Get An Estimate



10. Cups of Coffee Consumed in US in a week

Number of cups in the past week = number of cups per day × 7 (for 7 days per week).

The assumption treats each day equally, although calculation would be more precise if we mention that there are likely fewer cups consumed on the weekend, as people are not in the office and in general might feel less of a need to drink coffee.

Percent of the population that drinks coffee:
Assuming 300 million people in the U.S.
Assume that 20% are children that (we hope) do not drink coffee.
Also guess that another 20% of the population does not drink coffee at all (perhaps they prefer tea or other beverages, or just water).

Number of cups per day: here our guess is that of the remaining 60% of people, half drink 2 cups per day, a quarter drink 4 cups per day, and a quarter drink 1 cup per day.
This averages out to 2 × 0.5 + 4 × 0.25 + 1 × 0.25 = 2.25 cups per coffee drinker per day.

Therefore the calculation is = 60% × 2.25 × 300,000,000 = 405 million cups each day
405 million cups × 7 days per week = 2.84 billion cups per week
(you could round it to approximately 2.8 billion cups)

Note: the interviewer could then ask questions around how many cups are drunk at home or the office versus bought from a store, or other similar variations. The interviewer might also ask your thoughts on coffee trends and also how much revenue this would mean, etc.

11. Gmail Queries per second

Preliminary questions:

Do I get any additional background information?
- *Nope, that's it. I'm waiting for your number.*

Okay, I'm going to define a query as a Gmail operation. It could be a read, write or search operation. First, I'm going to estimate the number of Gmail users. Then, I'll estimate how often they use the service and how many operations they perform.
- *Go on*.

Starting from the top and working our way down:

Total population in the world.= *7 billion*
Internet usage in *developed countries* (likely to be) = *70 to 80 % range*.
However, it's much lower for *developing countries*
so on average for the whole world = 40 % of 7 billion people = 2. 8 billion people use the Internet.

The next number I need is the *percentage of people that use Gmail as their primary account*:

When I think about my friends, I'd say about 70 % use Gmail as their primary account. I know several that use other services. Given that my friends are more tech-savvy and American-centric than others, let's say worldwide only *20 %* use Gmail as their primary account.

That gives us, 0.2 *(7 billion people)* 0.4= 560 million

Gmail primary users:
Let's say the average Gmail user logs into their account about 4 times a week.
Each time they login, let's say they read on average 8 emails. Compose 2 emails and search for 1 email. Thus, the number of queries is:
= (4 Gmail logins per week) x ([8 read email operations per login] + [2 compose email operations per login] + [1 search email operation])
= 44 Gmail queries per week per user

Finally, to get number of Gmail queries per second, we do the following math:
= (560 million Gmail users) * (44 Gmail operations per week per user) * (1 week per 604,800 sec)
= 40.7k queries per second

There are approximately 41k Gmail queries per second.

12. iPhones sold in the US each year

For this question, I'd like to start with stating assumptions and then doing my top down calculations.

Starting with the assumptions:
- People in the United States (approx.) = 315 million
- People with cell phone = 90 %

- Cell phone users are locked into contracts. That is, they cannot get a new phone unless they pay a penalty to break a contract. The contract is normally two years, so cell phones get replaced once every 2 years.
- Each person buys on average 1 cell phone.
- Smartphones are about *60 percent* of new cell phone sales.
- Market share of the iPhone of the smartphone market = 40 %

Calculations:
The number of people looking to buy a new phone each year is
= (315 mn people of US)x(90% people having cell phone)x(50% will be buying a new phone this year)
= 142 million

The number of people that will buy the iPhone each year:
= (142 mn people that will be buying a new phone this year) x (60% will get a smartphone) x (40% of the smartphone buyers will buy an iPhone)
= 34 million

To conclude, 34 million iPhones are sold in the US each year.

13. Estimate how much it costs to run Flickr for a 20GB user.

Just a quick clarification, when you say 20GB user, are you saying that the user is using 20GB of storage, or they bought a plan, which allows them up to 20GB storage?
- *The latter.*

I assume you want me to calculate costs of actual usage. That is, if they signed up for a 20GB plan, and they use something less than 20GB. That's the number you want?
- *Yes.*

And you want me to estimate costs per year or per month?
- *Per month.*

Okay. Let's say that each Flickr user uploads 10 pictures per week. 1 picture is roughly 5 MB in size.
= 10 * 5 MB = 50 MB each week. =~ 2.5 GB of storage for the year.
But what we really need is the average cost at any given point in the year. So let's just take the midpoint, and the average user, in its first year has about 1.25 GB.

Let's talk about storage costs:
I'll estimate storage costs to be 10 cents per GB per month. (That's roughly how much Amazon S3 charges for storage)
We're storing 1.25GB per user, so cost = (1.25 GB) * ($0.10 per GB/month) = 12.5 cents/ month.
Per year = $1.50 per user.

From a bandwidth perspective, we usually show the optimized version of the photo. Let's assume optimization can reduce file sizes to 40%. So now the user sees, 5*0.4 = 2 MB version.

Assume each picture gets viewed on avg 10 times/ month, so 10x2 MB = 20 MB/mo. gets transferred.

We assumed that the user uploads 10 pictures per week = 520 photos per year.
Let's take the midpoint again, we get 260 photos. = 260 * 20 MB = 5.2GB per month.

Let's say the bandwidth costs = 12 cents per GB per month.
If we round our 5.2GB to 5GB, we get = 5GB* 12 cents = 60 cents per month per user.

14. Elevators needed for a 50-story building

I'm going to evaluate the number of elevators we need based on how many passengers we need to transport.

To start, I'll estimate how many people use the elevators and how long they're willing to wait. In the elevator industry, it's assumed that people are willing to wait 20 seconds for an elevator. It's also important to know how many people enter the building during peak times. I forgot to ask: is this building for office, retail, or residential use?
- *This is an office building.*

If the first part was all about elevator demand, this second part is about elevator capacity. I'd like to estimate:

- ➔ How many people can fit an elevator?
- ➔ What's the elevator's average speed?
- ➔ How long does an elevator stay open before closing and resuming travel?

It may also be important to know whether there are alternative transportation methods such as stairs or escalators. For the sake of simplicity, we'll set that aside.

Let's assume this is a 50 story office building and that each floor contains 120 employees. In total, that's 6,000 people in a 50 story office building.

People get to work between 8a.m.- 9a.m. The 6,000 people arrive uniformly during this 60 min. duration. That's 6000/ 60 = 100 people per minute.

We want people to wait no more than 20 seconds per elevator.
(100 people / 60 sec)*20 sec = 33 people, on average, wait 20 seconds.

To determine how many elevators are needed to transport 33 people once every 20 seconds, we need to determine the effective throughput of an elevator.

Let's assume there are 12 feet per floor and the elevator travels 20 ft /sec.
Also assume that in a 50 story building, the elevator stops 12 times on its way up, and each stop takes about 20 seconds.

Let's calculate the numbers:
There are 12*50 = 600 feet in this building.
At 20 ft/sec, it takes 600/ 20 = 30 seconds to get to the top.
However, there are 12 stops on the way up, so that takes 12stops * 20 sec per stop = 240 seconds.

On average, it takes = 240 sec at stops + 30 sec to reach top = 270 seconds to go up.
Let's say it takes half the time to make its way down in the morning = 270/2 = 135 seconds.
In total that's 270+135 sec = 6.75 minutes to go up and down.

Let's say each elevator can contain 10 people.
It'd take = 33 people every 20 sec (without waiting more than 20 sec) /10 people per elevator
= 3.3 elevators to transport the first batch of waiting passengers.

6.75 minutes will have elapsed before those first set of 3.3 elevators make their way back
As already assumed, every minute 100 people use the elevator.
So, in 6.75 minutes = (6.75*100) more people will come
Now for those people = (6.75*100)/10 = 67.5 more elevators will be required
So total 67.5+3.3 = **~ 71 elevators** are required

Sanity check:
Seattle's Columbia Tower has 83 floors, and they have 46 elevators

Comments: Candidate asked some good clarifying questions and demonstrated customer empathy. Nice approach of evaluating demand first, then capacity. A potential follow-up question is discussing elevator algorithms to service incoming requests.

15. Daily Global Revenue of Starbucks

Preliminary Questions

Do you mean Global Revenue for any given day?
-Yes, Calculate global revenue for starbucks for today.

Should I also include revenue generated by eatables sold along coffee too?
-Yes, consider eatables and coffee both for your analysis.

Should I consider special locations such as airports and offices?
- *We're looking for an overall ballpark figure.*

Assumptions:

- Since Starbucks is US-based, we consider that it generates 50% of its global revenue from its home country.
- Operating Hours are 6 AM to 10 PM.
- Average time to make a regular starbucks coffee is 5 minutes
- At maximum efficiency, 3 orders can be processed simultaneously in a normal store
- Average price of a normal coffee in Starbucks is $3
- For takeaways, people only take coffee.

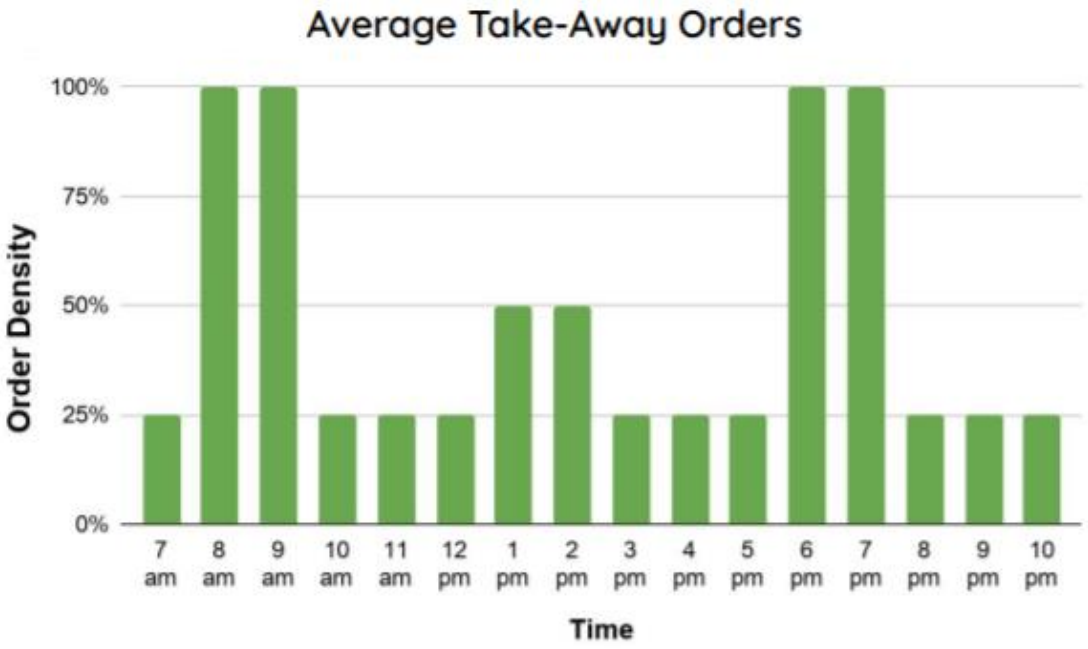
Overall Strategy:

Revenue Generated in USA = Average revenue of a store x # Stores in USA
Global Revenue of Starbucks = 2 x Revenue generated in USA

Average revenue of a Starbucks outlet in USA-
The revenue stream is divided into two types- Takeaway and In-Store Orders.

Takeaway Orders
Order density can be segmented as - Peak (100%), Medium (50%) and Low (25%). Peak time for takeaways are in the morning and evening as the working population grab coffee while commuting. Density is medium around lunch and low for the rest of the day.

During Peak Time, the store works at maximum efficiency; 3 coffees are prepared in 5 Minutes during peak time. Therefore, 36 Coffees/hour. Coffees sold in Medium Time - 18/hour Coffees sold in Low Time - 9/hour



Total Takeaway Orders in a day on average =
36 x 4 + 18 x 2 + 9 x 10 = 270 Coffee Orders/Day

Revenue Generated from Takeaways = No. of orders in a day x Average price of a takeaway order = 270 * 3 ~ 810$

In-Store Orders

For these, the peak time is usually after noon, when most people take breaks from their work and after office timings.

Total In-Store Orders in a day in a normal store = 36 x 6 + 18 x 5 + 9 x 5 ~ 350 Orders/Day
Average In-Store Order Cost = $5 (including eatables)

Revenue Generated from In-Store Orders =
No. of orders in a day x Average price of an in-store order
= 350 x 5 = ~ $1700
Total Revenue generated by an average store in a day = $810 + $1700 = $2510/day ~ $2500/day

Starbucks stores in the USA

Population of the United States = 300M = 30 Crores

Assuming that each Starbucks location caters to 2000 people in that locality,
Total # starbucks stores in USA = 30 crores/2000 = 15,000 Stores

US Revenue of Starbucks = Total Number of stores x Average Store Revenue
= 15,000 x $2500 ~ 37,500,000$

Global Daily Revenue = 2 x US Revenue = $75,000,000