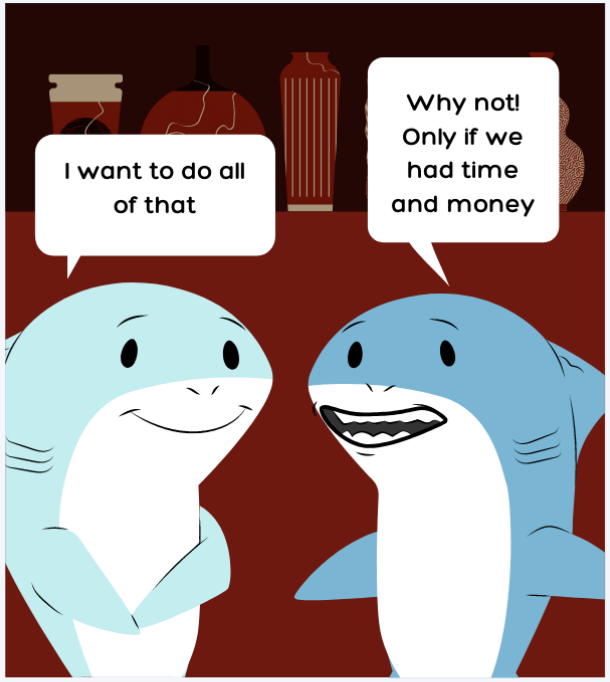
Tour Planner

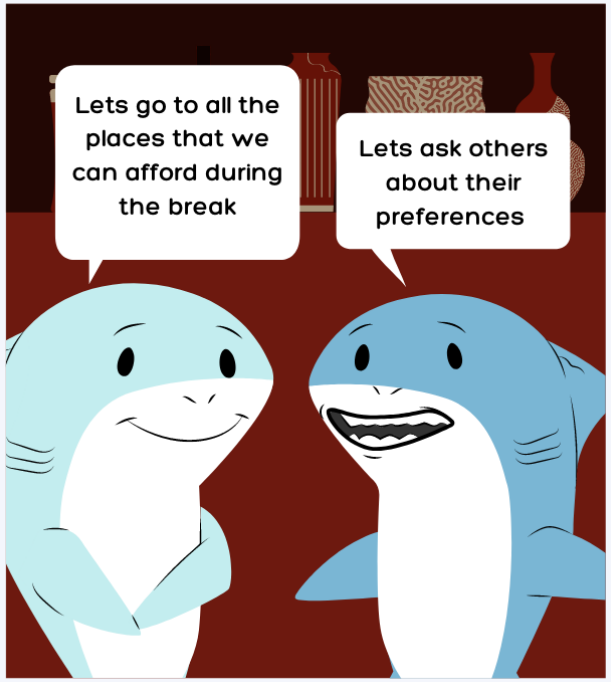
Group 4

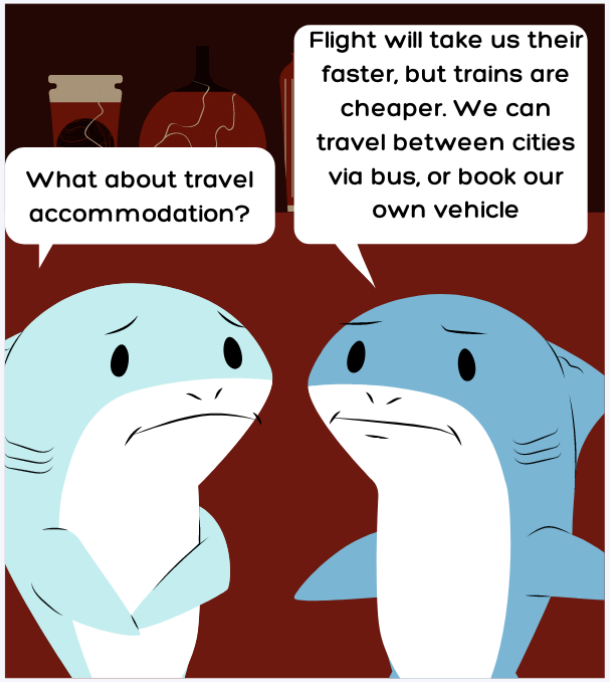
Aakash R | Prashant Ahirwar |Pratham Naik | Preetam Kumar

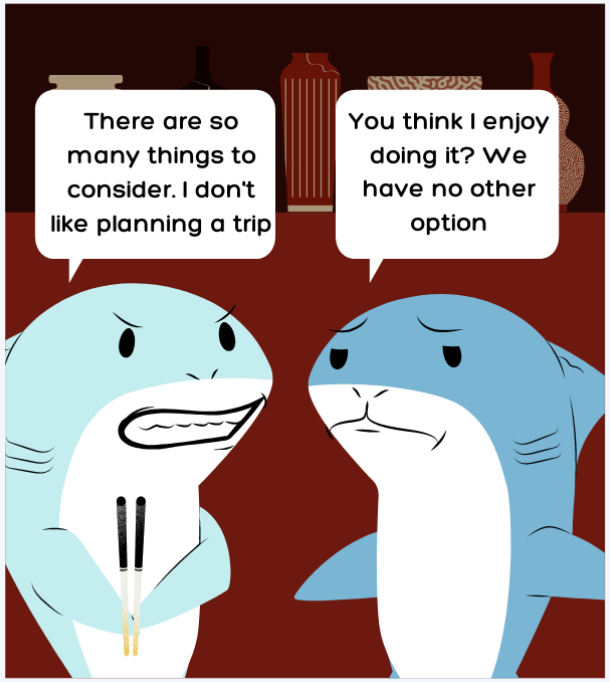

We have a term break, let's go for a trip to North India
Sure, what all places do you want to visit?


Anything you recommend?
We can go for treks, rafting, adventure sports, skiing. So much to choose from!

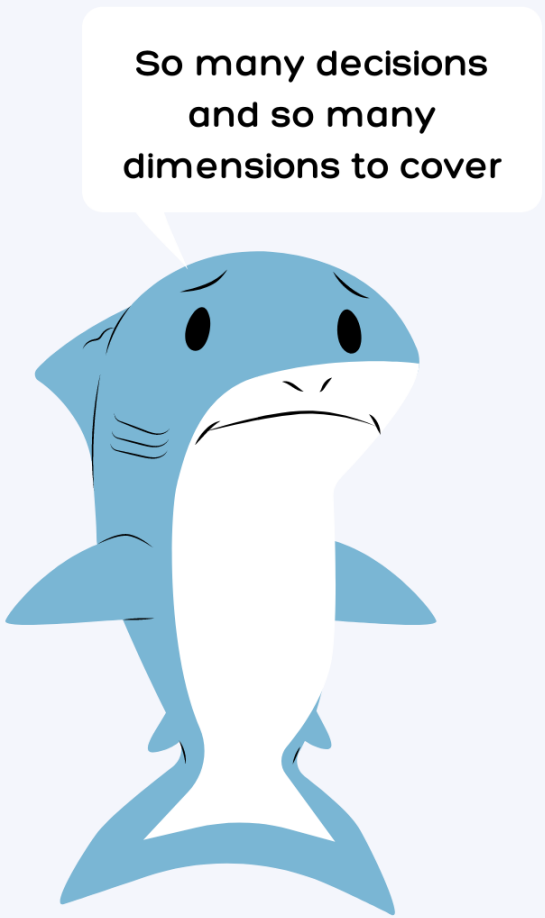

I want to do all of that
Why not! Only if we had time and money


Lets go to all the places that we can afford during the break
Lets ask others about their preferences


What about travel accommodation?
Flight will take us their faster, but trains are cheaper. We can travel between cities via bus, or book our own vehicle


There are so many things to consider. I don't like planning a trip
You think I enjoy doing it? We have no other option

Motivation



Want to travel as many places as possible

Have limited amount of time and money

Too many places for options

So many options for travelling

Want to cover all parts of the country

Want to cover different types of places

How do I know if I am not spending more than I need to spend

Problem:
Planning a small scale trip itself has too many decisions to make. For a small scale trip, one can manually evaluate multiple options and choose one. For a large scale trip, like and India wide tour, checking various options manually would not be possible and hence, one would usually end up planning a sub-optimal trip.

Solution:
We aim to develop a model which helps you with the major decisions involved in planning a trip.

Problem Description

Preetam wants to plan out an India wide tour for himself	
Objective	• I want to visit as many cities as possible in my tour from the list of cities I have in mind

Decisions	• I need to decide which cities will be covered as a part of this tour
	• For the cities I pick for the tour,
	  • I need to decide the sequence in which I will be visiting the cities
	  • I need to decide the number of days to stay in each city
	  • I need to decide the mode of travel for inter city travel for each pair of cities

Constraints	• I have limit on the amount of money I can spend on the trip • I have taken a vacation for the trip, so I have to complete the trip within 6 weeks • For each prospect city, based on the tourist spots/activities, I have decided the number of days that I should be spending there • I want to cover cities from all regions of the country • I want to cover different types of destinations like beaches, hill stations, historic locations, etc.

Illustrative Example Overview

Data for Illustrative Example

City ID	Hotel cost per day	Tourism expense	Recommended days to stay
0	0	0	0
1	1310	1000	3
2	1036	500	4
3	1387	4000	3

Parameter	Value
Budget	27000
Time	9
Daily expense	1100

Cost of Travelling Using Mode 1				
	0	1	2	3
0	0	5221	5704	5071
1	5862	0	5234	5146
2	5783	5065	0	5716
3	5752	5564	5435	0

Duration of Travelling Using Mode 1 (in hrs.)				
	0	1	2	3
0	0.00	2.94	2.76	2.12
1	2.36	0.00	2.50	2.12
2	2.23	2.61	0.00	2.90
3	2.41	2.72	2.99	0.00

Cost of Travelling Using Mode 2				
	0	1	2	3
0	0	2434	2571	2873
1	2435	0	2647	2420
2	2957	2752	0	2652
3	2516	2868	2164	0

Duration of Travelling Using Mode 2 (in hrs.)				
	0	1	2	3
0	0.00	10.19	10.31	10.19
1	10.91	0.00	10.44	10.49
2	10.22	10.86	0.00	10.11
3	10.11	10.68	10.96	0.00

Cost of Travelling Using Mode 3				
	0	1	2	3
0	0	4396	4380	4028
1	4111	0	4868	4472
2	4457	4731	0	4162
3	4406	4145	4968	0

Duration of Travelling Using Mode 3 (in hrs.)				
	0	1	2	3
0	0.00	8.10	8.73	8.59
1	8.59	0.00	8.82	8.57
2	8.81	8.09	0.00	8.04
3	8.57	8.50	8.58	0.00

- I want to visit as many cities as possible in my tour from the list of cities I have in mind

- I have limit on the amount of money I can spend on the trip
- I have taken a vacation for the trip, so I have to complete the trip within 6 weeks
- For each prospect city, based on the tourist spots/activities, I have decided the number of days that I should be spending there

A Non-trivial Problem

Possible Combination of Sequence and Mode of Transport = 675

Number of Cities to be Visited	1	2	3
Possible Option	3	6	6
Possible Mode Combination for each sequence	9	27	81
Total	27	162	486

Problem Formulation

Parameters

Parameter	Description
N	Number of prospect cities
B	Total budget for the trip
T	Available time for the trip
N_reg	Number of regions (North, East, West, South, Central)
N_cat	Number of categories (Hill station, beach, etc.)
$Cat_{i,c}$	1 if city i belongs to category c, else 0
$Reg_{i,r}$	1 if city i belongs to region r, else 0
Cat_min	Minimum number of cities to cover in each category
Reg_min	Minimum number of cities to cover in each region
$Days_i$	Recommended number of minimum days to stay in city i
$Hotel_i$	Hotel rent per night of stay in city i
$Path_{i,k,1,j}$	Price of k^{th} option for travel between city i and j
$Path_{i,k,2,j}$	Time of travel of k^{th} option for travel between city i and j
Daily_expense	Daily expenses for food and intra city travel (assumed to be constant)
$Tourism\_expense_i$	Expenses for tourist activities in city i (entry charges for tourist spots, adv. sports, etc.)
BigM	N+1

Sets

Set	Enumeration	Description
Nodes	0..N (0 for home city)	Set of prospect cities
Arcs	i in Nodes, j in Nodes : i ≠ j	Set of all possible combinations of 2 distinct prospect cities
Regions	1..N_reg	Set of al regions
Categories	1..N_cat	Set of all categories
Travel_options	1..9	Set of travel options (1..4 for flights, 5..8 for trains, 9 for road)
Dummy	1..2	1 for identifying price, 2 for identifying time

Variables

Variable		Description	Decisions
$X_{i,k,j}$	Binary	1 if option k is chosen for travel from city i to j, else 0	I need to decide which cities will be covered as a part of this tour I need to decide the mode of travel for inter city travel for each pair of cities
R_i	Integer >= 0	Sequence/rank of city i in the tour, 0 if not included in the tour	I need to decide the sequence in which I will be visiting the cities
$Time\_stay_i$	>=0	Number of days to stay in city i	I need to decide the number of days to stay in each city
Total_stay	>= 0	Total time for the tour	Just for reporting purposes
Total_expense	>= 0	Total expense for the tour	Just for reporting purposes

Objective Function

I want to visit as many cities as possible in my tour from the list of cities I have in mind

Maximize the number of cities that can be visited during the tour

maximize $(\sum_i . \sum_{j, j \neq i} . \sum_k x_{i,k,j} - 1)$

Number of travel options used to travel from city i to j

Number of cities visited from city i

Number of cities from which another city was visited
(= number of cities included in the tour)

Constraints

TSP Constraints (1/3)

No city can be visited from itself

$x_{i,k,i} = 0 \; \forall$ i in Nodes, k in Travel_options

Outdegree constraint for node 0: Exactly one city should be visited from city 0 (home city)

$$\underbrace{\sum_i . \sum_k . x_{i,k,0}}_{\text{Number of cities visited from city 0}} = 1$$

Outdegree constraint for other nodes: At-most one city should be visited from any city other than city 0

$$\underbrace{\sum_{j, j \neq i} . \sum_k . x_{i,k,j}}_{\text{Number of cities visited from city i}} \leq 1 \; \forall \text{ i in Nodes, } i \neq 0$$

Constraints

TSP Constraints (2/3)

Indegree constraint for other nodes: Each city other than city 0 can be visited from at-most one city

$$\underbrace{\sum_{i,i\neq j}\cdot \sum_{k}\cdot x_{i,k,j}}_{\text{Number of cities from which city j is visited}} \leq 1 \;\forall\; j \text{ in Nodes and } j \neq 0$$

Indegree = Outdegree: The number of cities visited from a given city should be the same as the number of cities form which the given city is visited

$$\underbrace{\sum_{i,i\neq m}\cdot \sum_{k}\cdot x_{i,k,m}}_{\text{Number of cities from which city m is visited}} = \underbrace{\sum_{j,j\neq m}\cdot \sum_{k}\cdot x_{m,k,j}}_{\text{Number of cities visited from city m}} \;\forall\; m \text{ in Nodes}$$

Constraints

TSP Constraints (3/3)

Rank of city j should be one more than the rank of city i if city j is visited from city i

$$R_j \geq R_i + 1 \underbrace{- \text{BigM}*(1 - \sum_k . x_{i,k,j})}\ \forall\ i,j \text{ in Arcs and } j > 0$$

This term is 0 if city j is visited from city i, else it is a large negative number making the constraint redundant

$$R_j \leq R_i + 1 \underbrace{+ \text{BigM}*(1 - \sum_k . x_{i,k,j})}\ \forall\ i,j \text{ in Arcs and } j > 0$$

This term is 0 if city j is visited from city i, else it is a large positive number making the constraint redundant

$$R_j \leq \text{BigM}*(\sum_i \sum_k . x_{i,k,j})\ \forall\ j \text{ in Nodes}$$

The maximum rank for any city should be less than or equal to the total number of cities, including home city

Rank for city 0 should be 1 because the tour starts from city 0

$$R_0 = 1$$

Constraints

Coverage Constraints

For each region, at least the minimum number of desired cities should be covered

$$\sum_i \sum_j \sum_k x_{i,k,j} * Cat_{j,c} \geq Cat\_\min \ \forall \ c \ \ in \ Categories$$

For each category, at least the minimum number of desired cities should be covered

$$\sum_i \sum_j \sum_k x_{i,k,j} * Reg_{j,r} \geq Reg\_\min \ \forall \ r \ \ in \ Regions$$

Constraints

Time Constraints

You should stay for the recommended number of days of stay in city i if city i is included in the tour

$$\text{Time\_stay}_i = \text{Days}_i * \underbrace{\sum_{j,j\neq i}. \sum_k. x_{i,k,j}}_{\text{1 if city i is included in the tour}} \quad \forall \text{ i in Nodes}$$

The tour should be completed within the available time

$$\text{Total\_stay} \leq T$$

Calculating the total duration of the tour

$$\text{Total\_stay} = \left(\frac{1}{24}\right) * \sum_i. \sum_{j,j\neq i}. \sum_k. (\text{path}_{i,k,2,j} * x_{i,k,j}) + \sum_i. \text{Time\_stay}_i$$

Constraints

Budget Constraints

Total expense of the tour should be within the budget

$$\text{Total\_expense} \leq B$$

Calculating the total expense of the trip

$$\text{Total\_expense} = \underbrace{\sum_i . \sum_{j, j \neq i} . \sum_k . (\text{Path}_{i,k,1,j} * x_{i,k,j})}_{\text{Travel expense}} + \underbrace{\sum_i . \text{Time\_stay}_i * \text{Hotel\_cost}_i}_{\text{Hotel expense}} + \underbrace{\text{Total\_stay} * \text{Daily\_expense}}_{\text{Daily expense}}$$

$$+ \underbrace{\sum_i . \sum_{j, j \neq i} . \sum_k . (\text{Tourism\_expensei} * x_{i,k,j})}_{\text{Tourism expense}}$$

Illustrative Example Solution

Data for Illustrative Example

City ID	Hotel cost per day	Tourism expense	Recommended days to stay
0	0	0	0
1	1310	1000	3
2	1036	500	4
3	1387	4000	3

Parameter	Value
Budget	27000
Time	9
Daily expense	1100

Cost of Travelling Using Mode[1] 1				
	0	1	2	3
0	0	5221	5704	5071
1	5862	0	5234	5146
2	5783	5065	0	5716
3	5752	5564	5435	0

Duration of Travelling Using Mode 1 (in hrs.)				
	0	1	2	3
0	0.00	2.94	2.76	2.12
1	2.36	0.00	2.50	2.12
2	2.23	2.61	0.00	2.90
3	2.41	2.72	2.99	0.00

Cost of Travelling Using Mode 2				
	0	1	2	3
0	0	2434	2571	2873
1	2435	0	2647	2420
2	2957	2752	0	2652
3	2516	2868	2164	0

Duration of Travelling Using Mode 2 (in hrs.)				
	0	1	2	3
0	0.00	10.19	10.31	10.19
1	10.91	0.00	10.44	10.49
2	10.22	10.86	0.00	10.11
3	10.11	10.68	10.96	0.00

Cost of Travelling Using Mode 3				
	0	1	2	3
0	0	4396	4380	4028
1	4111	0	4868	4472
2	4457	4731	0	4162
3	4406	4145	4968	0

Duration of Travelling Using Mode 3 (in hrs.)				
	0	1	2	3
0	0.00	8.10	8.73	8.59
1	8.59	0.00	8.82	8.57
2	8.81	8.09	0.00	8.04
3	8.57	8.50	8.58	0.00

- I want to visit as many cities as possible in my tour from the list of cities I have in mind
- I have limit on the amount of money I can spend on the trip
- I have taken a vacation for the trip, so I have to complete the trip within 6 weeks
- For each prospect city, based on the tourist spots/activities, I have decided the number of days that I should be spending there

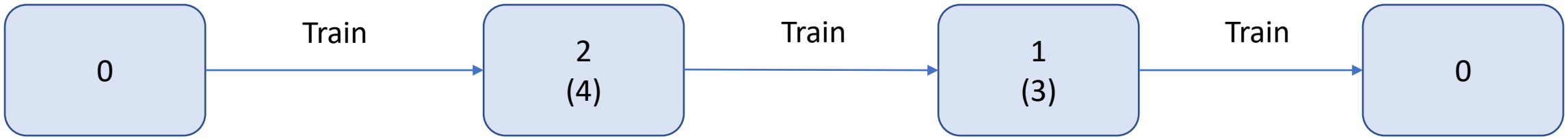
Footnote: [1]Here mode refers to airway, railway and roadway mode of transport

Solution

City index
(Time spent)

Requirements:	**Budget: Rs 27,000**	**Max Time: 9 Days**
Results:	**Expense: Rs 25,032**	**Total days: 8.3 Days**

Itinerary



Large Scale Example

Problem Specifications

30 cities from various states across the country (Home city: Delhi)

4 Regions:
North
South
East
West

4 Categories:
Hill station
Beach
Pilgrimage
Historical & Heritage

Top 4 flight results chosen for each pair of cities

Top 4 train results chosen for each pair of cities

Road transport cost calculated for each pair of cities by taking a price of Rs 10/km.

Flight and train ticket prices adjusted for inconvenient timings

For cases where 4 flights were not available, the corresponding slots were assigned a very high ticket price and time of travel

Data Collection

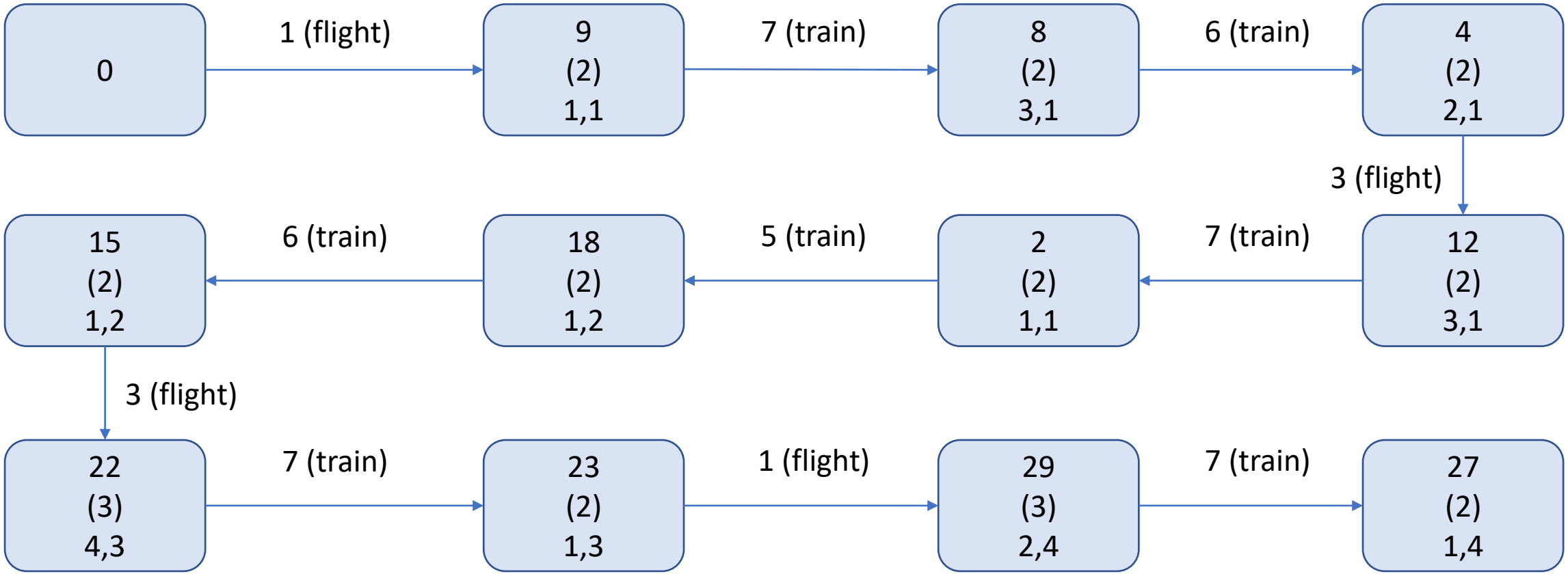
Roadways Data	• Manually collected longitude and latitude data for all prospect cities • Used Haversine formula to calculate bird-distance (refer to Appendix) • Used a multiplier to get roadways distance
Airways/Railways Data	• Scraped data from Goibibo.com and MakeMyTrip.com for airways and railways data respectively • Identified elements within the websites that can be used to access information for model parameters like departure time, arrival time, cost of ticket, flight/train number, etc. • Used selenium library with chrome-driver in Python
Hotel Data	• Hotel data was collected from MakeMyTrip.com by choosing price for the top result after applying appropriate filters (3 star hotel, 4+ user rating)
Daily expense	• This is a ballpark figure based on our experience
Tourism expense	• Since there was no direct source for this data, we have used random numbers for this parameter
Recommended number of stay days	• The data is based on personal experience and various internet sources
Adjustments	• For flights and trains arriving and departing at inconvenient times, we have added the cost of additional hotel stay in the price of the corresponding tickets. For example, if a flight arrives at 5 AM, you will have to take an additional day of hotel, so that additional cost of an extra day of stay in hotel is added to the price of the ticket. • For cities where direct options for flight and train were not available, the flights/trains to nearest airport/railway station were considered and road transport cost for travel from the nearest airport/railway station to the destination city was added to the price of the flight/train ticket price.

Solution

City index
(Time spent)
Category, Region

Requirements:	**Budget: Rs 1.05 Lakhs**	**Max Time: 30 Days**
Results:	**Expense: Rs 1.03 Lakhs**	**Total days: 24 Days**

Itinerary[1]



Footnote[1]: Itinerary represented is a simplified version of the actual result, the actual result represents info including but not limited to rail/airway name, departure/arrival time, etc.

Improvements

Appendix

Haversine Formula

Haversine:

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2\left(\frac{\Delta\lambda}{2}\right)$$

With:

$$c = 2 \cdot arctan2\left(\sqrt{a}, \sqrt{1-a}\right)$$

and:

$$d = R \cdot c$$

Where:

ϕ = latitude

λ = longitude

R = 6371 (Earth's mean radius in km)

Source: https://nathanrooy.github.io/posts/2016-09-07/haversine-with-python/