**OPEX ANALYTICS**

**CASE STUDY REPORT**

**Case 1.**

Objective: Given the supply chain network, find the minimum number of warehouses such that the following constraint is satisfied:

1. At least 80% of the demand by ton flows through the warehouses which are within 500 miles to the demand locations.

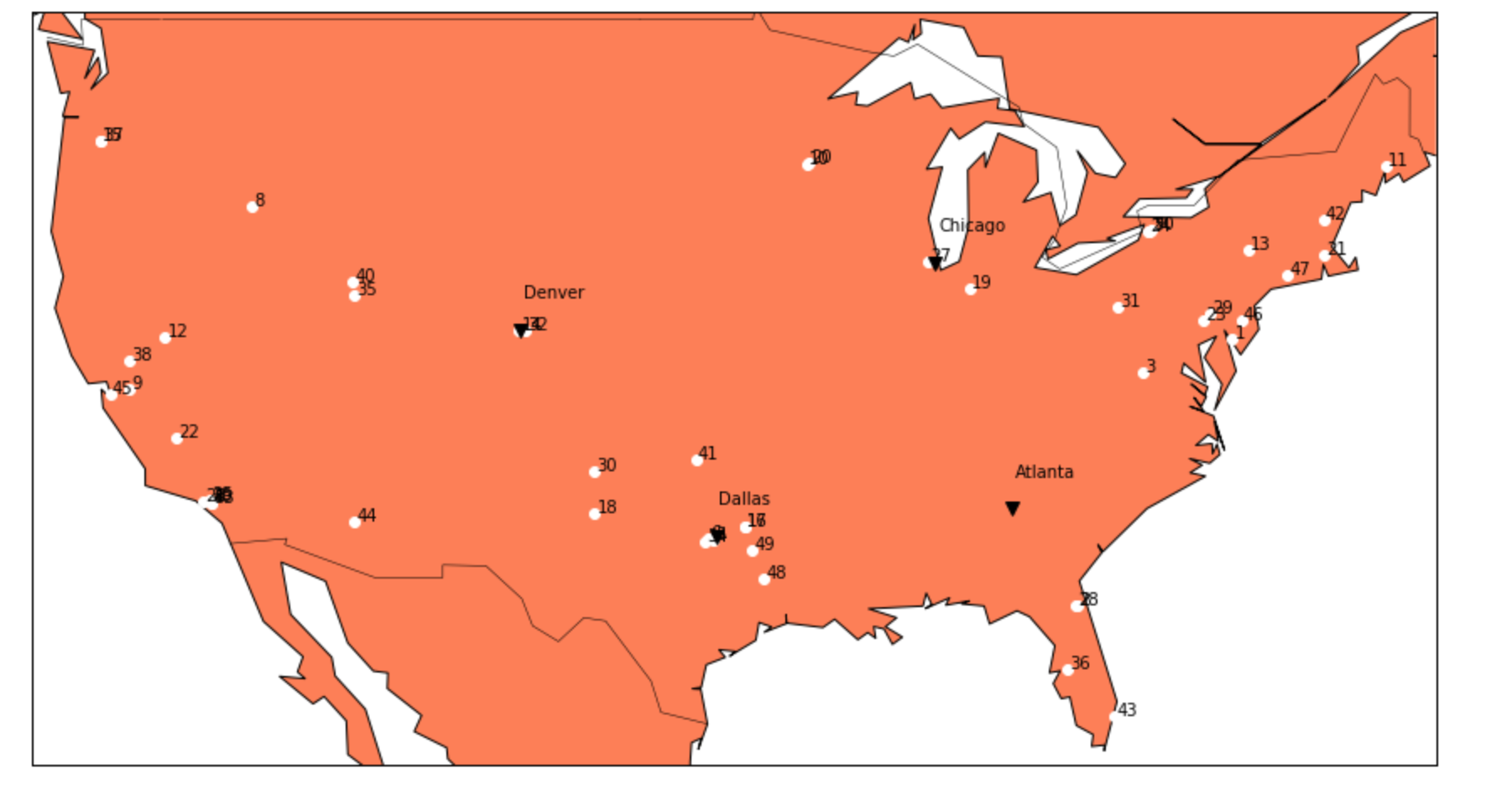
**Assumption**

The demand for various commodities at different plants was given for the years of 2012, 2013 and 2014. It is assumed that all the analysis is required to be done for the year of 2015 and the demand for the same has been calculated (approximated) using the average of previous years for each commodity at each demand point.

**Analysis**

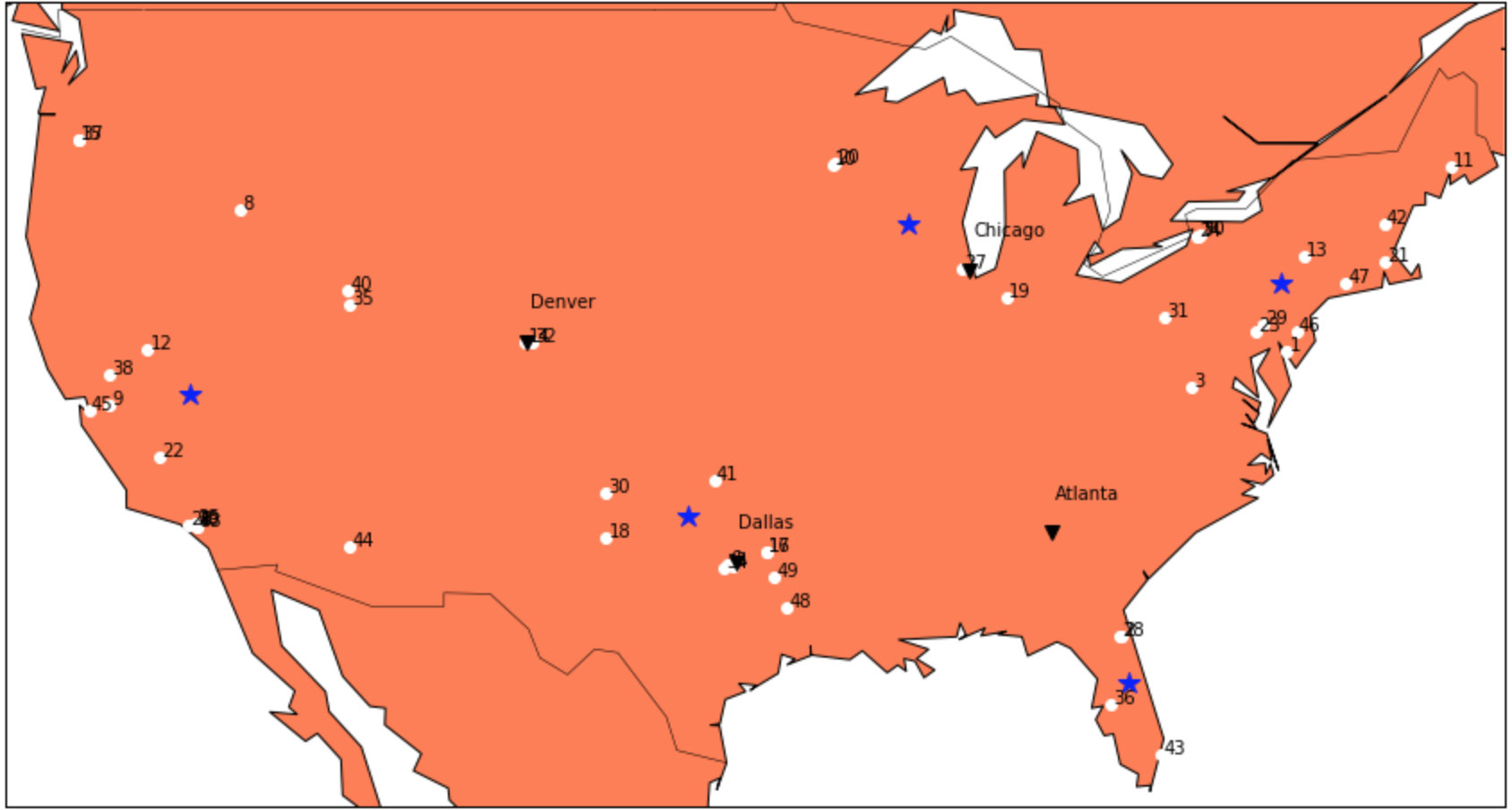
On analyzing the network, it became evident that the demand for product 1 (Clear) is the most in terms of quantity with 295488.7398 tons, followed by product 2 (Green) with 72580.14 tons, Red with 29797.477 tons, Blue with 11922.0901 tons and Grey with 5940.2655 tons.

Since the demand points are spread out across the United States, the analysis would become easier if the plants and demand points can be visualized. Hence, the following visualization has been created using python, with the white dots showing the demand locations and the black inverted triangles showing the plant locations.



*Figure 1: The supply chain network*

The map clearly shows that the demand points are concentrated in zones (4-5 zones: the northeast corridor, the midwest, the southwest, the west coast and the central zone). Hence, running an unsupervised clustering algorithm on the demand points such that the distance from warehouse to demand point criteria is met (<=500 miles) gives the following result: (the visualization)



*Figure 2: Supply chain network with prospective warehouse locations.*

The k-means clustering algorithm gives a set of clusters with a centroid (blue stars) depending on the features (distance and location, in this case) of the data points. The input for the number of clusters was chosen to be 5 (after roughly analyzing the demand points through the visualization).

Analyzing further, the aim is to minimize the number of warehouses. It is clear that the prospective warehouse(s) in Florida and Chicago aren’t close to much of the demand points. Hence, a demand supply analysis (summing up the demand of locations within 500 miles from the Florida and Chicago Warehouse and the same for the central America corridor) tells us that the first two can be merged and be put at demand location 8 (Boise, ID). A final check on the numbers and groups yields the following results:

Total load to be shipped out (tons) = 415728.75

80% of the total load (tons) = 332583

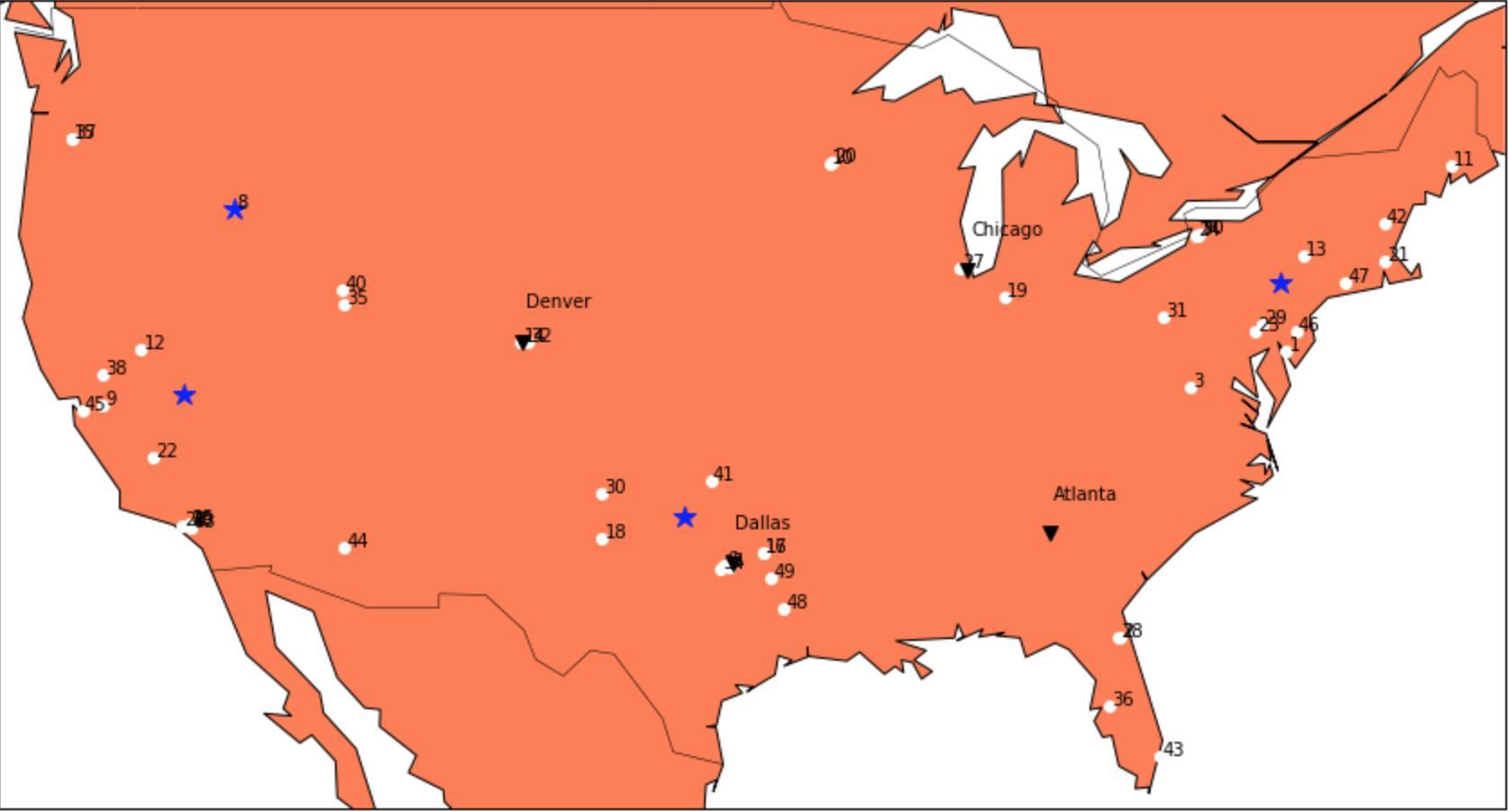
Groups/ Clusters as a result of clustering\*:

1. The northeast corridor with demand points (1, 3, 7, 11, 13, 21, 23, 24, 29, 31, 42, 46, 47, 50) with the warehouse placed at demand point 13
2. The south corridor with demand points (5, 6, 14, 16, 17, 18, 30, 32, 34, 41, 48, 49) with the warehouse placed at demand point 30.
3. The west coast corridor with demand points (4, 9, 12, 22, 25, 26, 33, 38, 39, 45) with the

warehouse placed at the demand point 22.

1. The central corridor with demand points (8, 15, 35, 37, 40) with the warehouse placed at

demand point 8.



*Figure 3: Proposed final warehouse locations (blue stars)*

Sum of load/ supply (tons) to each cluster:

The northeast corridor = 110303.89

The south corridor = 99149.20

The west-cost corridor = 88408.96

The central corridor = 64880.035

Summing all four up, the total load distributed = 362742.1 tons, which is more than 332583 tons. Since

the clustering and grouping analysis (in python notebook) made sure of the distance criteria, the

minimum warehouses required is equal to 4.

**Effect of transportation cost**

Background: Initially, (without the warehouses) the demand of each location was met by trucks shipping the goods directly from the plants to the respective demand locations (quarterly).

\*: It is assumed that the warehouses be placed at a demand point rather than at the centroid of the clustering itself. This made the calculations with respect to the transportation cost become easier.

By having warehouses, the transportation cost is bound to increase. The reason being, the increase in

the distance travelled by the goods (the shortest distance between two points is the straight-line connecting them). Hence, the increase in the transportation cost as found in the excel sheet.

Summary of transportation costs for the same quantity of load shipped, before and after warehousing:

**Before warehousing**:

Total trucks rented: 41671

Total transportation cost: $98,942,772.04

**After warehousing:**

Total trucks rented: 77891

Total Transportation cost: $124,514281.48

**Comments:**

The analysis clearly shows that building warehouses is not going to have a positive impact on cost aspect of the supply chain. Even if it is argued that the warehouse locations might have caused the inflation of prices, it is very evident that at least 4 warehouses are required to satisfy the given constraint. Also, the clustering algorithm (k-means) gives out a solution by minimizing the distance between the centroid and its neighboring points (hence, optimal in terms of location).

Also, the analysis does not include the initial investment required to build and maintain the four

warehouses which will further add up the cost.

**Conclusion:**

The idea of building warehouses in a radius of 500 miles from demand points such that 80% of the

demand is supplied through them increases transportation costs significantly.

The pros include better delivery service and organization.

The cons include significant increase in transportation costs and costs involved in building and

maintaining the warehouses.

**Case 2**

In order to save on transportation costs and improve service, it has been proposed that investments in

the production capabilities at each plant be made. These investments would allow each of the plants to

produce all products, albeit with a setup time and cost for switching from one product to another.

Objective: To model such a system and see how much impact it has on the supply chain in terms of cost and performance.

**Mathematical model:**

Min:

such that:

…..(1)

…….(2)

…(3)

Where, the decision variables are:

x(i,j,k): quantity(in tons) of commodity ‘j’ produced after set up change from commodity ‘i’ at plant ‘k’ in regular time.

x’(i,j,k): quantity (in tons) of commodity ‘j’ produced after set up change from commodity ‘i’ at plant ‘k’ in over-time.

y(c,a,b): amount of commodity ‘c’ shipped from plant ‘a’ to demand point ‘b’.

z(i,j): binary variable which becomes 1, if the decision variable x(i,j,k) or x’(i,j,k) is more than 0 for a particular (i,j).

The given values/data includes:

P denotes the production cost, OT denotes the production cost in over-time, D denotes the shipping cost and ST denotes the one-time set up cost of changing from producing commodity i to j.

Capk represents the capacity and Demand represents the demand of each demand location.

Constraints:

1. Constraint 1 is the capacity constraint, which says that the total production including over time should not exceed the capacity allotted for a commodity at a given plant.
2. The second constraint ensures that all demand is met.
3. The third constraint is the linking constraint which links the quantity of different commodities produced at the different plants and the quantities shipped out from there. (What is produced can only be shipped out!)

The rest are non-negativity and binary variable conditions. It is evident that the model is a MILP (Mixed Integer Linear Program).

Note on production and capacity constraint: Since the given data had information on the production rate of different plants and also the production capacities of different commodities, both couldn’t be incorporated in the model. (Eventually, both represent the same thing.). Also, although the demand has to be satisfied quarterly, since all quarters have the same demand, the total costs of fulfilling this would be the same as satisfying an annual demand once.

(% comment) The problem has been modelled and solved using MATLAB’s linprog solver. (The reason for using MATLAB is a software constraint. I don’t have a license for LINGO and gurobipy, hence the only tool which I could use (apart from CPLEX, which I’m unfamiliar with and need a license to use too) was MATLAB).

The software constraint forced the problem modelling to be much simpler/relaxed (LP). (With MATLAB, the binary constraint modelling couldn’t be done (to my knowledge and attempts to search for an alternative), because with MATLAB, the access is limited to the co-efficient of the decision variables and not the decision variables itself while modelling.)

The problem has been modelled and solved in such a way that it is scalable for all cases of investments (investing in just one plant, 2 plants, investing in 3 plants and investing in all 4 plants) by using the concept similar to Big M method. Rather than changing the input matrix every time, I’ve simply multiplied the cost of production of non-native commodities with a huge number at plants where the investment into making it as multi-functional hasn’t been made. (The final cost tells if the overprized commodity was produced, in which case, the problem would be declared infeasible.)

For example, in the three-plant investment case, the cost of making all other products per ton except ‘Clear’ at Chicago have been multiplied by a huge number (100000), prompting the model to not make any other products other than ‘Clear’ at Chicago.

**Logic behind choosing the plants for investment**

Analyzing the demands of each commodity along with the operational capabilities of each of the 4 plants, the following was observed:

|  |  |  |
| --- | --- | --- |
| Plant location | Demand for product (tons)/ annum | Production capacity (tons)/ annum |
| Chicago | 295488.7398 | 288000 + 144000 (R.T + O.T) |
| Dallas | 72580.14 | 144000 + 72000 |
| Denver | 29797.447 | 144000 + 72000 |
| Atlanta (combined 4 and 5) | 17862.35 | 144000 + 72000 |

R.T.: Regular time; O.T.: Over time

The table clearly illustrates that the demand for ‘Clear’, produced at Chicago is the most, resulting in the most production cost and transportation cost. Hence, investing at Chicago should be the least in priority since the demand for the product made at Chicago is the most.

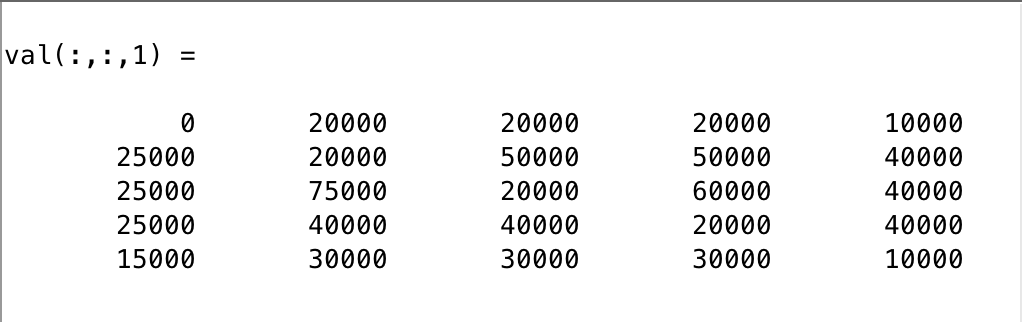
Also, Denver and Atlanta are under-utilized in terms of their production capacities. Hence, those two plants are the ones where an investment should be made first.

Using this logic, the following combinations were analyzed with the model:

1. Plant Atlanta is subject to investment and is capable of producing all commodities.
2. Plants Atlanta and Denver are subject to investment and are capable of producing all commodities.
3. Plants Dallas, Atlanta and Denver are subject to investment and are capable of producing all commodities.
4. All 4 plants are subject to investment and are capable of producing all commodities.

On looking at the solution for any case, say for example the case of investing in all 4 plants, it becomes clear that all four factories start producing all 5 commodities and there is no over-time production. If we look at the set-up cost (one-time cost per configuration change) separately, it becomes evident that the problem of not adding the binary constraint into the model can be easily solved (by making the conversions in the order that minimizes cost and the no. of days used for conversion).

Elaborating,



The above represents the set-up cost at plant 1 which initially produced ‘Clear’. The optimal way/order to minimize the set-up conversion cost is by solving it as a sequential tree problem (Game theory) or as minimum spanning tree. One such order for this plant is 1-4-2-5-3 which makes the conversion cost to be $130,000. (No need to worry about the no. of days aspect here because the plant finishes all its production within 150 days of regular time production exclusive of conversion days. More on this is on the excel sheet).

Hence, for the case of investing in all four plants, conversion days for plant 1 = 130000/ 5000 = 26 days (with the order of production 1-4-2-5-3). Using a similar approach, the order and costs for the remaining plants can be found as well, which turns out to be: $105,000 (2-1-5-3-4) for plant 2, $105,000 (3-1-5-4-2) for plant 3 and $115,000 (4-1-5-2-3) for plant 4. A similar approach can be adopted for all other combinations too.

**Comments:**

On analyzing all four cases, the following results were observed:

|  |  |  |  |
| --- | --- | --- | --- |
| **Case** | **Set-up cost (approximate)** | **Total cost (production + transportation)** | **Total cost after investment** |
| Invested only at Atlanta | $115,000 | $279, 356, 018.37 | $289, 356, 018.37 |
| Invested at Atlanta and Denver | $220,000 | $258, 129, 647.40 | $278, 129, 647.40 |
| Invested at Atlanta, Denver and Dallas | $325,000 | $245, 543, 481.57 | $275, 543, 481.57 |
| Invested at all 4 plants | $455,000 | $243, 740, 433.69 | $283, 740, 433.69 |

The above values are calculated using the model and appropriate parameters. These might not be the optimal values for each of the cases considered but should definitely be very close to the actual values as most of the relaxations in the model have been taken care of using analytical additions.

From the table, it is clear that investing in 3 plants (Atlanta, Dallas and Denver) turns out to be the least expensive. Even with adding the initial investment of $10,000,000 per plant, the third option saves ~$12 million a year! From the second year onwards, this investment would save ~42 million dollars as compared to the initial scenario!

Another benefit of this investment is that it gives a lot of scope for expansion in production. If invested in all three plants, the plants at Chicago and Atlanta finish their share of production in less than a year’s time, which gives lots of scope for producing more if the business expands.

**Conclusions:**

Investing in the three plants at Dallas, Denver and Atlanta so that the plants become capable of producing all 5 commodities, analytically proves to be the most profitable investment, yielding a saving of ~ $42 million per year, after the initial investment.

Although investing in all four plants yields more savings from the second year, the return of investment compared to investing in three plants seems to be less. (spending $10,000,000 to save ~2million more from the second year).

This investment also abides well along with the aim of the CEO to reduce transportation costs and enhance delivery quality.