

Analysis Documentation

Data Analyst: Ruby Smith

Created On: 1/22/2025

Updated On: 2/13/2025

- Access and import .csv files from MassageINC
- Create Analysis Report 'MassageINC Analysis Report'
- Start visually exploring data to better understand information.
- Begin preparation phase in Report
 - Includes: Purpose, Key Tasks, About Data, and Prepare Phase Conclusion
- Finish Preparation phase in Report
- **Begin cleaning "client information"**
 - Concatenate client names (First Name Last Name)
 - Updated Sales, Appointment, Cancellation, Payment Method names for all duplicated clients.
 - Overall: 10 duplicates removed.
- Create UniqueID for clients using DEC2HEX
- Check for misspelling and other discrepancies in filtered columns
- Remove Columns with insufficient data
 - "Source" Column - only 3 listed
 - "Alt mobile", "Alt phone", "landline", and "address 2"
- Update Birthdate column and create Age column.
- Change and update column names to: full_name, clientID, first_name, last_name, phone, address, city, state, zip, country, email, gender, birthday, age
- **Begin cleaning "sales" sheet**
 - Clean duplicate "FREE" in Sales that were created. 589 rows deleted. (25.44%)
 - Original row count: 2315
 - New row count: 1726
- Move Column D "Services" to end of columns, "Services" now Column J
 - Split "Service" column using a comma "," as the separator.
 - Update "Services" that are Tip Only as "No Service, Tip Charge"
 - Move Tips Calculated in "Services" into "Tips" Column
 - Clean columns created by splitting "Services" and Concatenate "Service" columns
 - Resplit Services to permanently keep separated, double check services.
- Clean "Payment Method" Column
 - Update methods that are "misspelled" or not categorized correctly
- Update "Duration" column and remove "min" after numbers
- VLOOKUP UniqueID for clients and add to "Sales" sheet
- Check for misspelling and other discrepancies in filtered columns
- Change and update Column Names to: purchase_date, appointment_date, client, clientID, duration_minutes, payment_method, sub_total, tips, total, therapist, services_1, services_2, services_3, services_4
- **Begin cleaning "appointment" sheet**

- Use VLOOKUP to include CLIENTID
 - Split “services” into 4 columns
 - Check for misspelling and other discrepancies in filtered columns
- Change and update column names to: appointment_date, appointment_time, client, clientID, price, therapist, service_1, service_2, service_3, service_4
- **Begin cleaning “cancellation” sheet**
 - Use VLOOKUP to include CLIENTID
 - Update “cancelled_by” names
 - Format date and time correctly
 - Split cancellation date/time into separate columns
- Change and update column names to: clientID, client_name, appointment_date, appointment_time, date_cancelled, time_cancelled, cancelled_by
- **Begin cleaning “payment Method” sheet**
 - Use VLOOKUP to include CLIENTID
- Filter “Free” payment method and take out discounted services.
 - Total = 589. Same as duplicates in sales.
 - Deleted 589 “free” duplicates.
 - Discrepancies between “sales” and “payment Method”. I believe issue is surrounding store credit/massage packages. Charges are shown in “payment method” but not “sales”. Used CLIENTID 2F3AB72E when checking.
- Import all .csv files to BigQuery to run SQL queries for analysis (except client_information)
- Begin analyzing services in appointment .csv and Sales .csv
 - Service data columns queried for types of services.
 - Percent difference in scheduled vs. sold.
 - Break down service types per year.
 - Create Total services per year with percent increase per year
 - Create estimate of services for rest of year.
- Analyze cancellations.csv
 - By month
 - By weekday
- Create list of clients with count of appt scheduled
 - Included gender ratio
- Create list of people who have created an account but have not booked an appointment
- Begin analyzing total sales in Sales .csv
 - Check Totals and separate years for rest of current year
 - Get and estimate future percent increases
 - Get Averages, Min, and Max, etc
- Create visuals through tableau using tables and aggregated data created with business data. <https://public.tableau.com/app/profile/ruby.smith/vizzes>
- Create slide show through google slides
<https://docs.google.com/presentation/d/1K7p4lILb3MHQdjC5P10W-fP6S9JpOPOP46QMhXaKn1o/edit?usp=sharing>
- Complete documentation and business analysis report.