

Università degli Studi di Udine
Dipartimento di Scienze matematiche, informatiche e fisiche
Internet of Things, Big Data & Machine Learning
A.A. 2022-23

Relazione del progetto di Basi di Dati

Gianluca Fabris 151884
Andrea Martin 147800
Andrea Brollo 147293
Alessio Stefan 152497

Indice:

- 1) Introduzione
- 2) Raccolta e analisi dei requisiti
 - 2.1) Glossario dei termini
 - 2.2) Ristrutturazione dei requisiti
 - 2.3) Requisiti Operazionali
- 3) Progettazione concettuale
 - 3.1) Modello Entità-Relazioni
 - 3.2) Lista Vincoli
 - 3.3) Tabella dei Volumi
 - 3.4) Caratteristiche delle operazioni
 - 3.5) Analisi delle Ridondanze
- 4) Progettazione logica
 - 4.1) Ristrutturazione del modello E-R
 - 4.2) Modello Relazionale
- 5) Progettazione fisica
 - 5.1) Analisi e scelta di opportuni indici
 - 5.2) Analisi forme normali
- 6) Implementazione
 - 6.1) Definizione di relazioni in SQL
 - 6.2) Popolamento della base di dati
 - 6.3) Definizione di alcuni trigger risultanti dalla progettazione
 - 6.4) Definizione di alcune query significative
- 7) Analisi dei dati in R
 - 7.1) Analisi delle ridondanze
 - 7.2) Analisi degli indici
 - 7.3) Esecuzione di alcune query significative
 - 7.4) Analisi dei dati presenti nella base di dati e altre analisi statistiche

1. Introduzione

Ci è stata consegnata la seguente tematica su cui creare il progetto di basi di dati:

Si voglia modellare il seguente insieme di informazioni riguardanti un sistema per la gestione degli animali di uno zoo.

- Ogni esemplare di animale ospitato è caratterizzato da un genere (ad esempio, zebra) e da un codice unico all'interno del genere di appartenenza. Per ogni esemplare, si memorizzano la data di arrivo nello zoo, il nome proprio, il sesso, il paese di provenienza e la data di nascita.
- Lo zoo è diviso in aree. Ogni area è identificata univocamente dal suo nome. In ogni area c'è un insieme di case, ognuna destinata ad un determinato genere di animali. Ogni casa è caratterizzata da un numero progressivo (1, 2, 3, ...) che la identifica univocamente all'interno dell'area di appartenenza. Ogni casa contiene un insieme di gabbie, ciascuna contraddistinta da una lettera (A, B, C, ...), che la identifica univocamente all'interno della casa. Ogni gabbia contiene un solo esemplare. Ogni casa ha un addetto che pulisce tutte le gabbie in essa presenti in un determinato giorno della settimana. Uno stesso addetto può avere l'incarico di pulire più case (in giorni diversi della settimana).
- Gli animali sono sottoposti periodicamente a controllo veterinario. In un controllo, un veterinario rileva il peso degli esemplari, diagnostica eventuali malattie e prescrive il tipo di dieta da seguire.

Si definisca uno schema Entità-Relazioni che descriva il contenuto informativo del sistema, illustrando con chiarezza le eventuali assunzioni fatte. Lo schema dovrà essere completato con attributi ragionevoli per ciascuna entità (identificando le possibili chiavi) e relazione. Vanno specificati accuratamente i vincoli di cardinalità e partecipazione di ciascuna relazione. Si definiscano anche eventuali regole di derivazione e/o vincoli di integrità necessari per codificare alcuni dei requisiti attesi del sistema.

2. Raccolta e analisi dei requisiti

In questa fase abbiamo analizzato il testo della consegna e i requisiti per la nostra base di dati, definendo un glossario dei termini e in seguito creando alcune tabelle per ristrutturare i requisiti. Infine abbiamo specificato qualche operazione utile per il nostro Zoo.

Nel nostro caso il testo della consegna era già stato pre filtrato e riorganizzato prima del nostro intervento e quindi parte del nostro lavoro è stata svolta al solo scopo di seguire il processo di raccolta e analisi dei requisiti che è stato introdotto a lezione.

Abbiamo prestato particolare attenzione al dominio di riferimento, affinché la base di dati che andremo a modellare possa rispecchiare le caratteristiche reali che potrebbero essere richieste da uno Zoo.

2.1. Glossario dei termini

Il glossario evidenzia ogni termine importante, a cosa si riferisce e con quali altri termini è collegato.

termine	descrizione	relazioni
esemplare di animale	esemplare di una specie di animale (singolo animale)	genere, veterinario
genere	specie di animale	esemplare di animale
area	parte dello zoo	casa

casa	parte dell'area	area, gabbia
gabbia	parte della casa dove c'è un esemplare di animale	casa, addetto
addetto	colui che pulisce le gabbie	gabbia
veterinario	colui che esegue il controllo veterinario agli esemplari di animale	esemplare, controllo veterinario
controllo veterinario	visita per il controllo di salute dell'esemplare di animale eseguita dal veterinario	veterinario

2.2. Ristrutturazione dei requisiti

In questa fase abbiamo evidenziato i requisiti da soddisfare per ciascun elemento all'interno della nostra base di dati, creando la seguente tabella; inoltre abbiamo implementato anche una leggenda per poter evidenziare con maggiore semplicità le varie entità, attributi, chiavi e relazioni.

Leggenda
grassetto = entità <i>corsivo</i> = attributo <u>sottolineato</u> = chiave evidenziato in grigio = relazione

Frase di natura generale
Si voglia modellare il seguente insieme di informazioni riguardanti un sistema per la gestione degli esemplari di animali di uno zoo
Frase relative agli esemplari di animali
Ogni esemplare di animale ospitato è caratterizzato da un genere (ad esempio: zebra) e da un <u>codice unico</u> all'interno del genere di appartenenza. Per ogni esemplare di animale , si memorizzano la <i>data di arrivo</i> nello zoo, il <i>nome proprio</i> , il sesso, il <i>paese di provenienza</i> e la <i>data di nascita</i> .
Frase relative alle aree
Lo zoo è diviso in aree . Ogni area è identificata univocamente dal suo <i>nome</i> In ogni area c'è un insieme di case , ognuna destinata ad un determinato genere di animali
Frase relative alle case
In ogni area c'è un insieme di case ...

<p>Ogni casa è caratterizzata da un <u>numero progressivo (1, 2, 3, ...)</u> che la identifica univocamente all'interno dell'<u>area</u> di appartenenza</p> <p>Ogni casa <u>contiene</u> un insieme di gabbie...</p> <p>Ogni casa ha un addetto che <u>pulisce</u> tutte le gabbie...</p>
Frase relative alle gabbie
<p>Ogni casa <u>contiene</u> un insieme di gabbie, ciascuna contraddistinta da una <u>lettera (A, B, C, ...)</u>, che la identifica univocamente all'interno della casa.</p> <p>Ogni gabbia <u>contiene</u> un solo esemplare di animale.</p> <p>Ogni casa ha un addetto che <u>pulisce</u> tutte le gabbie.</p> <p>Uno stesso addetto può avere l'<u>incarico di pulire</u> più case (in diversi <i>giorni della settimana</i>)</p>
Frase relative agli addetti
<p>Ogni casa ha un addetto che <u>pulisce</u> tutte le gabbie in essa presenti in un determinato <i>giorno della settimana</i></p> <p>Uno stesso addetto può avere l'<u>incarico di pulire</u> più case (in diversi <i>giorni della settimana</i>)</p>
Frase relative al controllo veterinario
<p>Gli esemplari di animali sono <u>sottoposti</u> periodicamente a controllo veterinario</p> <p>In un controllo veterinario, un veterinario...</p>
Frase relative al veterinario
<p>In un controllo veterinario, un veterinario rileva il <i>peso</i> degli esemplari di animali, diagnostica eventuali <i>malattie</i> e prescrive il tipo di <i>dieta</i> da seguire</p>

2.3. Requisiti Operazionali

Per finire, abbiamo definito le seguenti operazioni idealmente utili per la realtà che dobbiamo realizzare:

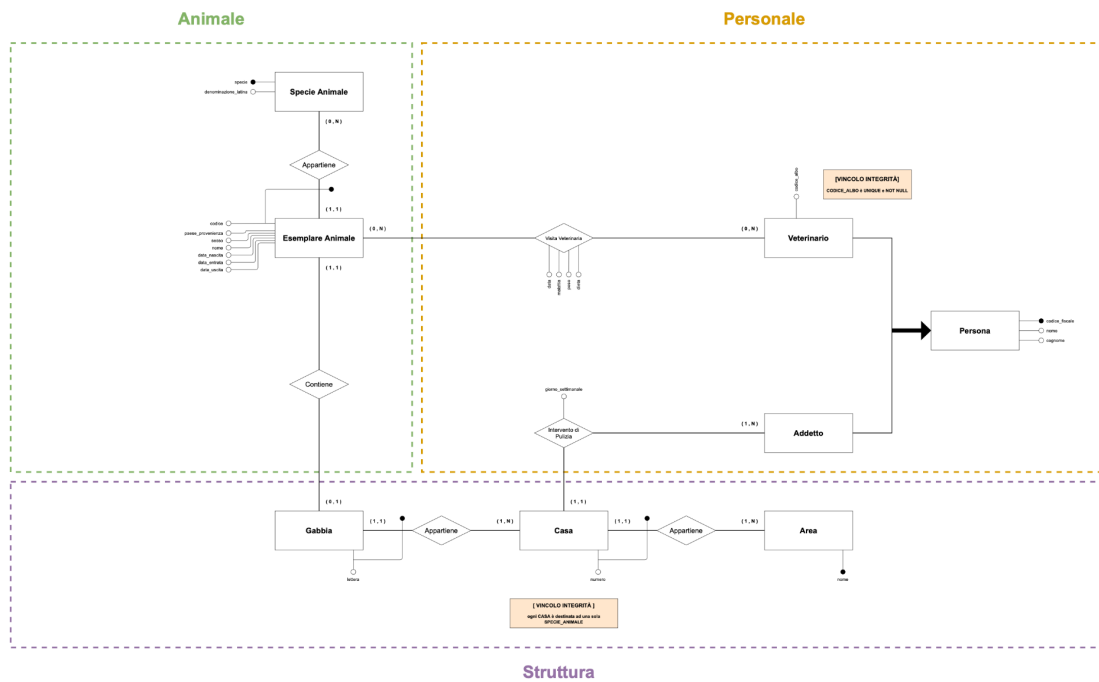
- **Operazione 1:** ultimo controllo veterinario di un esemplare di animale
- **Operazione 2:** inserimento di un controllo veterinario di un esemplare di animale
- **Operazione 3:** lista esemplari di animale (attuali) presenti nello zoo, divisi per aree
- **Operazione 4:** dato un esemplare di animale visualizzare la gabbia, casa, area
- **Operazione 5:** eliminazione / inserimento dell'esemplare di animale
- **Operazione 6:** visualizzazione delle gabbie vuote
- **Operazione 7:** esemplari di animale con ultimo controllo veterinario oltre un mese fa
- **Operazione 8:** media controlli veterinari veterinarie per ogni specie
- **Operazione 9:** numero di controlli veterinari effettuati su una esemplare di animale rispetto la sua specie

3. Progettazione concettuale

In questa fase abbiamo costruito un modello concettuale di tipo entità-relazione, definendo i vincoli di integrità e le regole di derivazione associati, ed in seguito abbiamo effettuato l'analisi delle ridondanze su di esso.

3.1. Modello Entità-Relazioni

Dall'analisi del testo abbiamo prodotto il seguente schema entità-relazioni.



Nel modello seguente abbiamo introdotto la storicizzazione delle Visite Veterinarie e il numero_visite_veterinarie come attributo derivato. Queste scelte sono state fatte durante una seconda analisi di controllo effettuata sul modello che era appena stato progettato, in maniera tale da migliorarlo.

- La relazione Intervento di Pulizia lega le entità Addetto e Casa ed ha come vincoli (1, 1) e (1, N) perché per ogni Casa è assegnato un solo Addetto (per l'attività di pulizia), ma nell'arco della settimana un Addetto può pulire più di una Casa
- La relazione Visita Veterinaria Corrente lega direttamente il Veterinario e l'Esemplare di Animale aggiungendo anche degli attributi della relazione che sono funzionali all'identificazione della Visita Veterinaria. La relazione ha come vincoli (1, N) e (0, N) perché ogni Animale ha ricevuto almeno una Visita Veterinaria (quella mensile) mentre un Veterinario può aver visitato più di un Esemplare di Animale però potrebbero essere tutte Visite Veterinarie delle mensilità precedenti (quindi righe appartenenti alla storicizzazione)
- La relazione Visita Veterinaria Storico lega direttamente il Veterinario e l'Esemplare di Animale e rappresenta la storicizzazione

3.2. Lista Vincoli

3.2.1. Regole di Derivazione

- Il **numero di visite veterinarie** effettuate per ogni esemplare di animale è derivabile contando il numero di interventi veterinari registrati per il singolo animale selezionato;

3.2.2. Vincoli di Integrità

Sono emersi dall'analisi dei requisiti i seguenti requisiti:

- Le case possono contenere una sola specie di animale ognuna.
- In un'area è possibile avere più case con specie diverse
- Una casa è pulita solo 1 volta, alla settimana, da 1 solo addetto.
- Una gabbia può tenere solo 1 animale
- Un veterinario per operare nello zoo, deve essere registrato.

Inoltre, abbiamo definito i seguenti vincoli:

- Se la gabbia è vuota continuo a pulirla
- Non eliminare la specie di animale anche se non possiedi più esemplari di animali appartenenti alla specie di animale di interesse.
- Una casa può apparire una volta sola nella tabella Interventi di Pulizia.
- In Interventi di Pulizia la coppia Area-Casa deve essere unica (chiave candidata).
- Quando un animale non è più presente nello zoo, il campo Data di Uscita cambia da null alla data di uscita.
- Quando un animale viene trasferito, la gabbia diventa vuota.
- Se la Visita Veterinaria viene effettuata da un veterinario esterno, esso non viene registrato come veterinario bensì viene utilizzato un record fittizio (veterinario esterno)
- Quando termina il contratto di lavoro con una Persona registrata nel database, essa non viene eliminata bensì mantenuta per poter evidenziare un precedente contratto in caso di riassunzione

3.3. Tabella dei Volumi

nero = entità

blu = relazioni

Entità / Relazione	Quantità
Specie Animale	50
Esemplare Animale	1000

Area	6
Casa	40
Gabbia	200
Veterinario	100
Addetto	20
Esemplare Animale appartiene a Specie Animale	1000
Esemplare Animale è Contenuto in una Gabbia	200
Gabbia Appartiene a Casa	200
Casa Appartiene ad Area	40
Casa ha Intervento di Pulizia con Addetto	40
Esemplare di animale ha Visita Veterinaria Storico con Veterinario	100000
Esemplare di animale ha Visita Veterinaria Corrente con Veterinario	200

Entità:

- abbiamo ipotizzato di avere, nel nostro Zoo, circa un 1000 Esemplari Animale (inclusi i non presenti in struttura) con una quantità di Esemplari Animale presenti in struttura pari a 200 unità circa (in ogni giornata lavorativa) e all'incirca 50 Specie Animale (non necessariamente nello stesso istante e non divise equamente)
- si suppone che il nostro Zoo sia suddiviso in 6 aree costituite da 7 case ognuna e che per ogni casa siano disponibili 5 gabbie (per un totale di 40 case e 200 gabbie).
- per gestire tutte le attività necessarie al mantenimento dello Zoo e degli Esemplari Animale abbiamo pensato che il personale Veterinario sarà composto circa da 5 dottori in contemporanea per arrivare ad un massimo di 20 dottori totali contando anche i non presenti in struttura, per le pulizie invece ipotizziamo che siano presenti 40 Addetti (uguale al numero di Case da tenere pulite) che, contando i non attivi possono raggiungere un massimo di 100 all'interno della base di dati.

Relazioni:

- ipotizzando di non gestire la storicizzazione della programmazione degli Interventi di Pulizia si prende il numero di Case come numero massimo di Interventi (40 Interventi).
- le Visite Veterinarie Storico invece sono all'incirca 100.000, tenendo conto del fatto che in questo caso viene gestita la storicizzazione delle visite passate (in maniera tale da avere un minimo di tracciabilità), mentre Visite Veterinarie Correnti sono in totale 200 (come il numero di esemplari di animale)
- l'appartenenza di uno specifico Esemplare Animale ad una Specie Animale conta 1000 occorrenze

- il contenimento di un Esemplare Animale in una specifica Gabbia è presente 200 volte (numero di esemplari correnti)
- L'appartenenza ad una Casa per ogni Gabbia occorre 200 volte (= numero di esemplari e di gabbie)
- Il contenimento di una Casa all'interno di una specifica Area è presente 40 volte

3.4. Caratteristiche delle operazioni

Operazione	Tipo	Frequenza
Operazione 1	I	200 al mese
Operazione 2	I	200 al mese
Operazione 3	I	10 al giorno
Operazione 4	I	2500 al giorno
Operazione 5	I	2 a settimana
Operazione 6	I	2 a settimana
Operazione 7	I	1 al giorno
Operazione 8	I	100 al mese
Operazione 9	I	100 al mese

Le operazioni dalla 1 alla 9 sono di tipo "I" (**Interattive**), perché molto frequenti. Non ci sono di tipo "B" (**Batch**), ovvero eseguite poco frequentemente.

Nota: non potendo conoscere concretamente le dinamiche interne di uno Zoo, abbiamo ipotizzato le frequenze delle operazioni cercando di attribuire valori il più possibile credibili.

Per la frequenza delle operazioni 1 e 2 abbiamo pensato di 200 accessi al mese (tenendo in considerazione l'idea di effettuare 1 visita al mese per ogni esemplare di animale presente nello zoo).

Per l'operazione 3 sono stati ipotizzati 10 accessi al giorno, per poter ottenere una copia del database utilizzabile da vari chioschi all'interno della struttura e all'amministrazione.

Nell'operazione 4, a causa dei possibili continui accessi da parte dei chioschi (per visitatori) e dei veterinari che devono effettuare le visite veterinarie, abbiamo ipotizzato che ci siano 2500 accessi al giorno.

L'operazione 5 e l'operazione 6 ipoteticamente potrebbero avere 2 accessi a settimana per motivi di vendita, acquisto, morte di un esemplare di animale.

L'operazione 7 quasi sicuramente verrà eseguita con una frequenza di 1 volta al giorno dal pc dell'amministrazione per poter informare i veterinari degli esemplari di animali a cui è stata effettuata l'ultima visita veterinaria oltre un mese fa.

L'operazione 8 e la 9 si ipotizza che abbiano una frequenza pari a 100 accessi al mese.

Queste operazioni, idealmente, vengono chiamate per vari motivi, tra cui:

- prima dell'acquisto di un nuovo esemplare di animale (poter acquistare un esemplare di animale più economico dal punto di vista delle spese mediche),
- in caso di crisi posso decidere quali esemplari di animali vendere ad altri zoo (per poter risparmiare sulle spese)
- per poter attribuire un grado di salute ad ogni esemplare di animale (identificare gli esemplari più deboli in maniera tale da poter notificare ai veterinari la situazione)
- per poter attribuire un grado di affidabilità ad ogni veterinario
- a scopo statistico

3.5. Analisi delle Ridondanze

In questa sezione abbiamo svolto un'accurata analisi delle ridondanze rispetto ad alcune operazioni precedentemente definite. Per questa fase ci siamo serviti delle tabelle degli accessi, analizzandole in presenza di ridondanza e in assenza di essa.

3.5.1 Analisi delle ridondanze per l'inserimento di una nuova visita veterinaria (Operazione 2)

L'operazione consiste nello spostamento dei dati da Visita Veterinaria Corrente a Visita Veterinaria Storico, in modo da gestire la storicizzazione con l'inserimento di una nuova visita per animale. Dopo aver aggiornato lo storico, viene modificata l'entità Visita Veterinaria Corrente in modo da contenere i valori dell'ultima visita effettuata.

3.5.1.a Operazione in presenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
Visita Veterinaria Corrente	R	1	R
Visita Veterinaria Storico	R	1	W
Visita Veterinaria Corrente	R	1	W
Esemplare Animale	E	1	R
Esemplare Animale	E	1	W

In caso di ridondanza si deve anche aggiornare l'attributo numero_visite_veterinarie nell'entità Esemplare Animale (ultime due righe della tabella riguardano l'aggiornamento dello stesso)

Costo singola operazione: **8R**

Costo Totale: **costo singolo * 200 = 1.600R**

L'operazione richiede **1.600** accessi in lettura.

3.5.1.b Operazione in assenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
----------	------	---------	--------------

Visita Veterinaria Corrente	R	1	R
Visita Veterinaria Storico	R	1	W
Visita Veterinaria Corrente	R	1	W

Costo singola operazione: **5R**

Costo Totale: **costo singolo * 200 = 1.000R**

L'operazione richiede **1.000** accessi in lettura.

3.5.2 Analisi delle ridondanze per l'attributo derivato numero_visite_veterinarie (operazione 8)

L'operazione consiste nella lettura dell'attributo calcolato numero_visite_veterinarie in Specie Animale in relazione agli animali presenti.

3.5.2.a Operazione in presenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
Esemplare_Animale	E	1.000	R
Appartiene	R	1.000	R
Specie_Animale	E	50	R

In caso di ridondanza si deve anche aggiornare l'attributo numero_visite_veterinarie nell'entità Esemplare Animale (ultime due righe della tabella riguardano l'aggiornamento dello stesso)

Costo singola operazione: **2050R**

Costo Totale: **costo singolo * 100 = 205.000R**

L'operazione richiede **205.000** accessi in lettura.

3.5.2.b Operazione in assenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
Visita Veterinaria Storico	R	100.000	R
Visita Veterinaria Corrente	R	200	R
Esemplare Animale	E	1.000	R
Appartiene	R	1.000	R
Specie Animale	E	50	R

Costo singola operazione: **102.250R**

Costo Totale: **costo singolo * 100 = 10.225.000R**

In assenza di *numero_visite_veterinarie* nell'entità Esemplare Animale, per dover calcolare lo stesso dato avremo bisogno di **10.225.000** accessi in lettura.

3.5.3 Analisi delle ridondanze per l'attributo derivato *numero_visite_veterinarie* (operazione 9)

L'operazione consiste nella lettura dell'attributo calcolato *numero_visite_veterinarie* in Esemplare Animale

3.5.3.a Operazione in presenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
Esemplare Animale	E	21	R
Appartiene	R	20	R
Specie Animale	E	1	R

Costo singola operazione: **42R**

Costo Totale: **costo singolo * 100 = 4.200R**

L'operazione richiede **4.200** accessi in lettura.

3.5.3.b Operazione in assenza di ridondanza

Concetto	Tipo	Accessi	Tipo Accessi
Visita Veterinario Storico	R	100	R
Visita Veterinaria Corrente	R	1	R
Esemplare Animale	E	21	R
Appartiene	R	20	R
Specie Animale	E	1	R

Costo singola operazione: **143R**

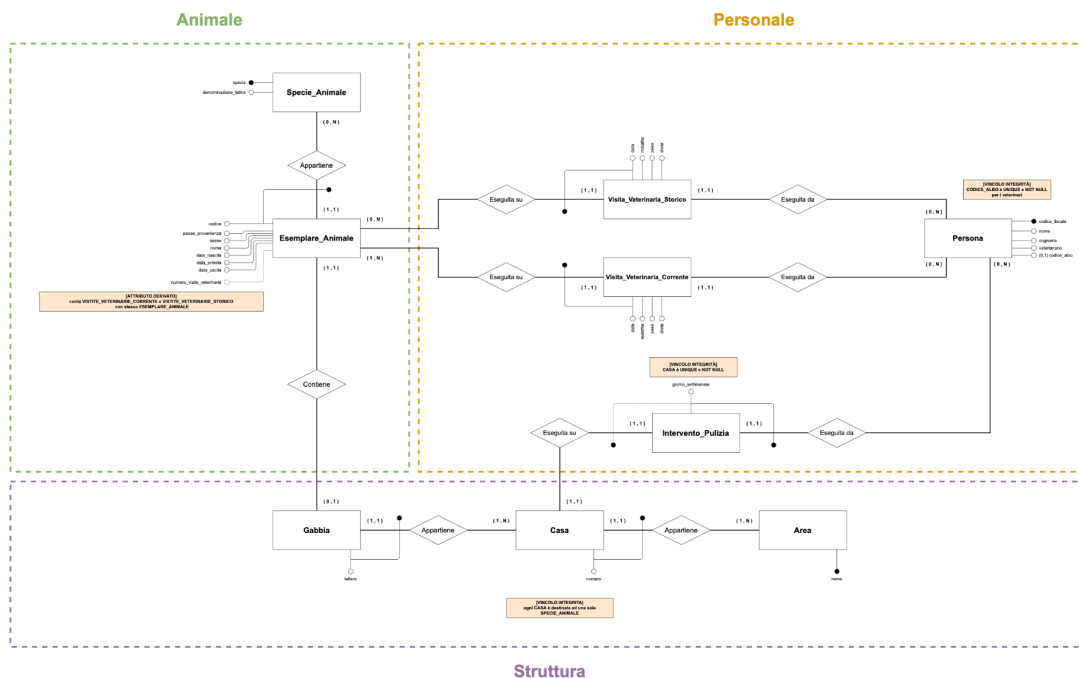
Costo Totale: **costo singolo * 100 = 143.000R**

In assenza di *numero_visite_veterinarie* nell'entità Esemplare Animale, per dover calcolare lo stesso dato avremo bisogno di **143.000** accessi in lettura.

4. Progettazione logica

In questa fase abbiamo effettuato la ristrutturazione del modello concettuale e basandoci su esso abbiamo costruito un modello logico (relazionale).

4.1. Ristrutturazione del modello ER



Ci sono tre modi per ristrutturare la generalizzazione nel modello iniziale (punto 3.1):

1. Eliminare le due entità figlio (Veterinario e Addetto) per poi aggiungere all'entità padre (Persona) gli attributi dei figli.

Le entità risultanti avrebbero una partecipazione opzionale (0,1) ed un attributo tipo_persona che indica se è Veterinario (V) o Addetto (A).

Le relazioni 'Visita Veterinaria' e 'Intervento di Pulizia' con l'entità 'Persona' hanno partecipazione facoltativa, perché:

- Solo i Veterinari possono avere la relazione 'Visita Veterinaria'
- Solo gli Addetti possono avere la relazione 'Intervento di Pulizia'

2. Tenere le due entità figlio (Veterinario e Addetto) ed eliminare l'entità padre (Persona).

Veterinario e Addetto acquisiscono gli attributi di Persona.

La chiave primaria di entrambe le entità è la chiave primaria di Persona.

Le due entità avranno entrambe le stesse relazioni con le altre entità che aveva il padre.

3. Tenere sia l'entità Persona che le entità Veterinario e Addetto.

Creando una relazione tra ciascun figlio e il padre; con vincoli (0,1) dalla parte del padre, e (1,1) dalla parte del figlio.

I figli avranno come chiave primaria la chiave esterna che punta al padre (cioè, in questo caso, il Veterinario e l'Addetto avrebbero come chiave primaria il campo codice_fiscale di Persona, che sarebbe anche la chiave esterna).

È stato deciso di adottare l'opzione 1, in quanto gli attributi specifici per Veterinario o Addetto erano alquanto limitati, e questa opzione ci permette di semplificare la base di dati.

L'opzione **2** comporta un incremento della complessità delle due tabelle figlie senza nessun beneficio, per tanto è stata scartata.

L'opzione **3** è utile quando le operazioni si riferiscono ad occorrenze e attributi di Veterinario (Addetto) o di Persona, e quindi fanno distinzioni tra entità figlio e padre. Tuttavia, non abbiamo riscontrato queste distinzioni.

Inoltre sono state reificate le relazioni "Visita Veterinaria Corrente", "Visita Veterinaria Storica" e "Intervento di Pulizia" in entità, in quanto contengono attributi significativi. Inoltre sono state aggiunte le relative relazioni per collegare le nuove entità al resto della base di dati.

Nella entità "Intervento di Pulizia" abbiamo individuato una possibile chiave candidata, oltre la chiave primaria, utile per l'identificazione della n-upla.

4.2. Modello Relazionale

Leggenda:

- Nome della tabella = **Testo colorato in arancione**
- Chiave primaria (quindi anche not null e unique) = sottolineatura
- Not null = **grassetto**
- Unique = *italic*
- Chiavi Esterne = CE: attributo → Tabella

Specie_Animale (specie, denominazione_latina)

attributo derivato: media_visite_veterinarie → media Esemplare_Animale con stessa Specie_Animale

Esemplare_Animale (specie, codice, paese_provenienza, sesso, nome, data_nascita, data_entrata, data_uscita, numero_visite_veterinarie)

CE: specie → Specie_Animale(specie)

attributo derivato: numero_visite_veterinarie → conta Visita_Veterinaria_Corrente e Visita_Veterinaria_Storico con stesso Esemplare_Animale

Area (nome)

Casa (numero, area)

CE: area → Area(nome)

Gabbia (lettera, casa, area, specie, animale)

CE: casa, area → Casa(numero, area)

CE: animale, specie → Esemplare_Animale(codice, specie)

Persona (codice_fiscale, nome, cognome, veterinario, codice_albo)

codice_albo se veterinario=1

Intervento_Pulizia (addetto, giorno_settimanale, **area**, **casa**)

CE: casa, area → Casa(numero, area)

CE: addetto → Persona(codice_fiscale)

Visita_Veterinaria_Storico (data, animale, specie, malattia, peso, dieta, veterinario)

CE: animale, specie → Esemplare_Animale(codice, specie)

CE: veterinario → Persona(codice_fiscale)

Visita_Veterinaria_Corrente (data_animale_specie, malattia, peso, dieta, veterinario)

CE: animale, specie → Esemplare_Animale(codice, specie)

CE: veterinario → Persona(codice_fiscale)

5. Progettazione fisica

Il linguaggio che abbiamo utilizzato per la progettazione fisica del database è PostgreSQL (che è stato illustrato dai docenti del corso).

Nella parte seguente è stata sviluppata tutta la sezione relativa alla scelta degli indici e definizione delle relazioni in linguaggio SQL.

5.1. Analisi e scelta di opportuni indici

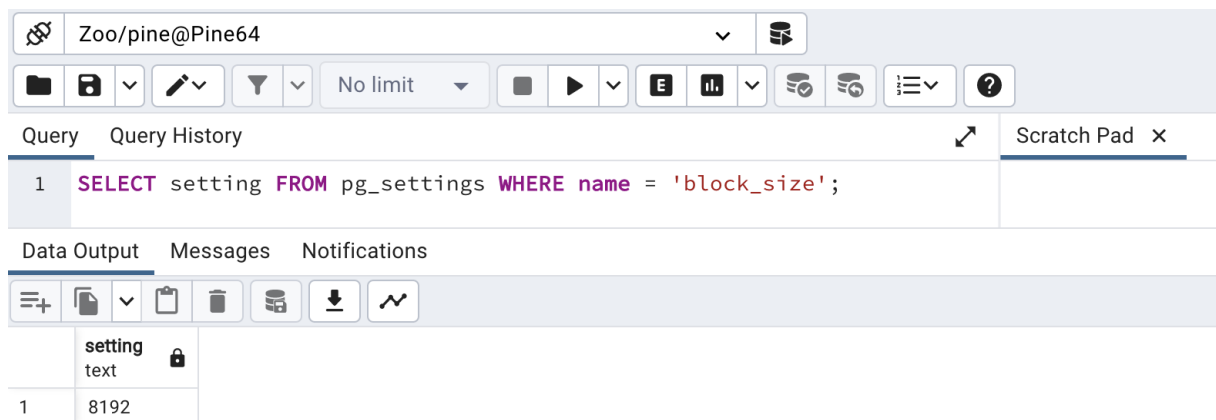
Gli indici sono strutture dati ausiliarie che garantiscono un rapido accesso ai dati in corrispondenza di determinate operazioni.

Oltre agli indici primari, cioè gli indici specificati rispetto al campo chiave di file di record fisicamente ordinati rispetto a tale campo (ordering key field), che vengono generati automaticamente dal DBMS, abbiamo ritenuto opportuno implementare i seguenti indici secondari, i quali possono essere definiti su qualsiasi campo.

Tutti i nostri indici sono stati scelti per il medesimo motivo, ovvero che le tabelle in cui sono stati introdotti possiedono una densità di righe molto elevata e in un utilizzo comune della nostra base di dati ci sarebbero un numero veramente elevato di richieste.

Il codice SQL per creare gli indici elencati nel file indici.sql.

Per recuperare il valore di B è bastato andare a vedere (tramite la seguente query) il valore impostato di default nel nostro database.



The screenshot shows a PostgreSQL client interface. At the top, there's a connection bar with 'Zoo/pine@Pine64'. Below it is a toolbar with various icons. The main area is divided into 'Query' and 'Query History' tabs. The 'Query' tab is active, showing a single query: `1 SELECT setting FROM pg_settings WHERE name = 'block_size';`. To the right of the query is a 'Scratch Pad' tab. Below the query editor are tabs for 'Data Output', 'Messages', and 'Notifications'. The 'Data Output' tab is active, displaying a table with one row and one column. The column is labeled 'setting' with a text icon and a lock icon. The row contains the value '8192'.

	setting
1	8192

Inoltre, in PostgreSQL 15, il puntatore al blocco (ovvero il puntatore utilizzato per accedere ai blocchi di dati all'interno del database, identificato con la lettera P) è di 8 byte.

5.1.1 Calcolo del costo computazionale sull'indice data_visita_storico_indice

5.1.1.a Costo in assenza dell'indice

$B = 8.192$

$r = 100.000$

$R = 8+4+50+100+9+1.000+16 = 1.187$

$bfr = B/R = 8.192/1.187 = 6$ per blocco
 $nb = r/bfr = 100.000/6 = 16.667$ blocchi
 $\log_2 nb = \log_2 16.667 = 14$ accessi

5.1.1.b Costo in presenza dell'indice

$V = 8$

$P = 8$

$nb_1 = nb = 16.667$

$R_i = 8+6 = 16$

$f_o = bfr_i = B/R_i = 8.192/16 = 512$

$t = \log_{f_o} nb_1 = \log_{512} 16.667 = 2$ livelli indici

$nb_2 = nb_1/f_o = 16.667/512 = 33$ blocchi

$nb_3 = nb_2/f_o = 33/512 = 1$ blocco

$t+1 = 2+1 = 3$ accessi

5.1.2 Calcolo del costo computazionale sull'indice data_visita_corrente_indice

5.1.2.a Costo in assenza dell'indice

$B = 8.192$

$r = 200$

$R = 8+4+50+100+9+1.000+16 = 1.187$

$bfr = B/R = 8.192/1.187 = 6$ per blocco

$nb = r/bfr = 200/6 = 34$ blocchi

$\log_2 nb = \log_2 34 = 5$ accessi

5.1.2.b Costo in presenza dell'indice

$V = 8$

$P = 8$

$nb_1 = nb = 34$

$R_i = 8+6 = 16$

$f_o = bfr_i = B/R_i = 8.192/16 = 512$

$t = \log_{f_o} nb_1 = \log_{512} 34 = 1$ livello indici

$nb_2 = nb_1/f_o = 34/512 = 1$ blocco

$t+1 = 1+1 = 2$ accessi

5.1.3 Calcolo del costo computazionale sugli indici data_entrata_indice e data_uscita_indice

5.1.3.a Costo in assenza dell'indice

$B = 8.192$

$r = 1.000$

$R = 4+50+1+50+8+8+8+50+4 = 183$

$bfr = B/R = 8.192/183 = 44$ per blocco

$nb = r/bfr = 1.000/44 = 23$ blocchi

$\log_2 nb = \log_2 23 = 4,5$ accessi

5.1.3.b Costo in presenza dell'indice

$V = 8$

$P = 8$

$nb_1 = nb = 23$

$R_i = 8+6 = 14$

$f_o = bfr_i = B/R_i = 8.192/16 = 512$

$t = \log_{f_o} nb_1 = \log_{512} 23 = 1$ livelli indici

$nb_2 = nb_1/f_o = 23/512 = 1$ blocco

$t+1 = 1+1 = 2$ accessi

5.2. Analisi forme normali

Verificando le tabelle e le dipendenze funzionali, si può evincere che sono sempre rispettate le seguenti proprietà:

1. Tutti gli attributi rappresentano informazioni atomiche
2. Tutte le righe di una tabella hanno lo stesso numero di colonne
3. Tutti i valori di una colonna sono dello stesso tipo
4. Ogni riga è diversa da tutte le altre
5. L'ordine delle righe non è importante
6. Tutti gli attributi non chiave dipendono dall'intera chiave
7. Tutti gli attributi non chiave dipendono solo dalla chiave e non da altri attributi

Le prime 5 proprietà garantiscono la **prima** forma normale.

Aggiungendo la sesta proprietà si raggiunge la **seconda** forma normale.

Aggiungendo la settima serve a garantire la **terza** forma normale.

6. Implementazione

I file contenenti i dati utilizzati sono resi disponibili nella cartella SQL.

6.1. Definizione di relazioni in SQL

Abbiamo creato, mediante l'utilizzo del linguaggio SQL, il file create.sql che permette all'utente di creare lo schema e le tabelle della nostra base di dati.

Questo file definisce la loro struttura, i relativi vincoli e le varie relazioni a livello fisico.

6.2. Popolamento della base di dati

Per popolare la base di dati, realizzata in questo progetto, abbiamo utilizzato www.mockaroo.com (che era stato proposto in sede di lezione dal docente) per la maggior parte delle tabelle presenti nel nostro database.

Si vuole evidenziare il fatto che mockaroo riesce quasi sempre a popolare le tabelle con dati verosimili; nei casi in cui i dati non sono coerenti con l'aspettativa bisogna intervenire con programmi esterni.

Un esempio concreto può essere quello del codice fiscale (come nel nostro caso); il codice generato da mockaroo non è conforme ai dati dell'utente registrato e quindi per avere un codice fiscale "valido" avremmo potuto intervenire con uno script python che presi nome, cognome e data di nascita genera un codice fiscale legato ai dati reali dell'utente.

```
codice_fiscale.py U •
codice_fiscale.py > ...
1  import datetime
2
3  vocali = ['A', 'E', 'I', 'O', 'U']
4
5  def getFirstPart(stringa):
6      ref = []
7      conta_lettere_stringa = 0
8      upper_string = stringa.upper()
9      for lettera in upper_string:
10         if conta_lettere_stringa < 3:
11             if lettera not in vocali:
12                 ref.append(lettera)
13                 conta_lettere_stringa = conta_lettere_stringa + 1
14         if len(ref):
15             for lettera in upper_string:
16                 if conta_lettere_stringa < 3:
17                     if lettera in vocali:
18                         ref.append(lettera)
19                         conta_lettere_stringa = conta_lettere_stringa + 1
20             return ''.join(ref)
21
22
23  def getCode(firstname, lastname, date):
24      res = getFirstPart(lastname) + getFirstPart(firstname) + date.strftime("%y") + 'B' + date.strftime("%d") + 'G888D'
25      return res
26
```

Questa sarebbe stata una possibile implementazione di una funzione python che può generare un codice fiscale più valido (non completamente corretto per mancanza del luogo di nascita).

Per la chiave esterna delle gabbie e per le visite veterinarie corrente (e visite veterinarie storico inserite tramite visite veterinarie corrente) abbiamo usato R per creare più facilmente i dati necessari. Abbiamo anche usato R per inserire i dati sul server Postgres.

Il codice R che ha popolato le nostre tabelle lo si può trovare direttamente nella cartella del progetto, nella cartella SQL.

6.3. Definizione di alcuni trigger risultanti dalla progettazione

Nel file trigger.sql, abbiamo definito alcuni trigger utili per la nostra base di dati:

- Inserimento di un nuovo esemplare di animale in una gabbia comporta al controllo che ogni casa destinata ad una sola specie come vincolo d'integrità **controlla_inserimento_animale_gabbia**
- Inserimento di una nuova visita veterinaria comporta lo spostamento della visita veterinaria corrente in visita veterinaria storico e calcolo dell'attributo derivato numero_visite_veterinarie **controlla_inserimento_visita_veterinaria**
- Eliminazione da Esemplare_Animale è gestita mettendo la data di uscita e spostando l'ultima visita veterinaria corrente in visita veterinaria storico **controlla_eliminazione_esemplare**

Per impedire l'accesso in scrittura all'utente a Esemplare_Animale.numero_visite_veterinarie e a Visita_Veterinaria_Storico, ma permettere ai trigger di eseguire le operazioni in automatico, abbiamo modificato le autorizzazioni dell'utente.

6.4. Definizione di alcune query significative

Il file query.sql contiene le query per le 9 operazioni del punto 2.3 e alcune operazioni che abbiamo ritenuto significative.

Di seguito sono presentate alcune possibili query che potrebbero essere effettuate sul database.

```
query1.sql
1  --operazione: visualizzare lo storico delle visite veterinarie di un esemplare di animale
2  select data, animale, Visita_Veterinaria_Storico.specie, malattia, peso, dieta, veterinario
3  from Visita_Veterinaria_Storico, Esemplare_Animale
4  where animale = codice and Visita_Veterinaria_Storico.specie = Esemplare_Animale.specie and codice = ###INPUT###
5  union
6  select data, animale, Visita_Veterinaria_Corrente.specie, malattia, peso, dieta, veterinario
7  from Visita_Veterinaria_Corrente, Esemplare_Animale
8  where animale = codice and Visita_Veterinaria_Corrente.specie = Esemplare_Animale.specie and codice = ###INPUT###;
9
```

```
quer2.sql
1  --operazione: visualizzare i turni dei interventi di pulizia dei rispettivi giorni della settimana, addetti, case e aree
2  select giorno_settimanale, codice_fiscale, Persona.nome, cognome, numero as casa, Area.nome as area
3  from Intervento_Pulizia, Persona, Casa, Area
4  where addetto = codice_fiscale and casa = numero and Intervento_Pulizia.area = Casa.area and Casa.area = Area.nome;
5
```

La stringa ###INPUT### è un placeholder per un dato che l'utente può scegliere.

7. Analisi dei dati in R

In questa parte vengono presentati i risultati delle analisi che sono state effettuate con R sui dati generati nella sezione precedente per il database.

Il file in R usato per fare queste analisi è reso disponibile nella cartella R.

7.1. Analisi delle ridondanze

Abbiamo eseguito l'analisi delle ridondanze con la seguente formula ottenuta dai calcoli della tabella delle ridondanze $-3x + 100200y + 101z > 0$, al fine di valutare meglio la convenienza delle stesse (grafici ottenuti utilizzando la libreria plotly di R).

Le variabili x, y e z sono le frequenze delle operazioni, rispettivamente operazione 2, operazione 8 e operazione 9.

Data la complessità della formula, l'abbiamo divisa e analizzata in 3 parti atomiche:

- **R1:** la funzione risolta rispetto a **Z**; risulta sempre conveniente avere la ridondanza
- **R2:** la funzione risolta rispetto a **Y**; risulta quasi sempre conveniente avere la ridondanza, l'unica parte in cui non conviene è per piccoli valori di X (sotto 30)
- **R3:** la funzione risolta rispetto a **X**; risulta sempre conveniente avere la ridondanza, ma non è sensata in quanto la ridondanza non verrebbe mai aggiornata

In seguito siamo riusciti a creare un grafico (**R4**) con tutte le variabili contemporaneamente su un campo più ristretto di dati.

Questo grafico conferma quanto detto in precedenza: risulta quasi sempre conveniente avere la ridondanza, l'unica parte in cui non conviene è per piccoli valori di Z (sotto 30).

Figura R1

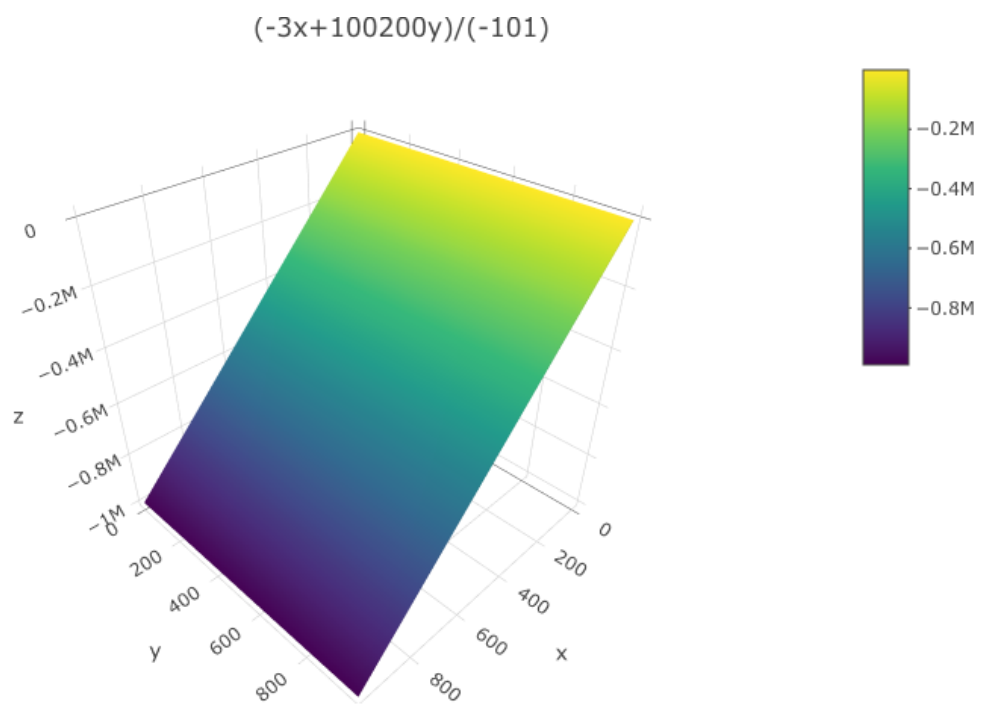


Figura R2

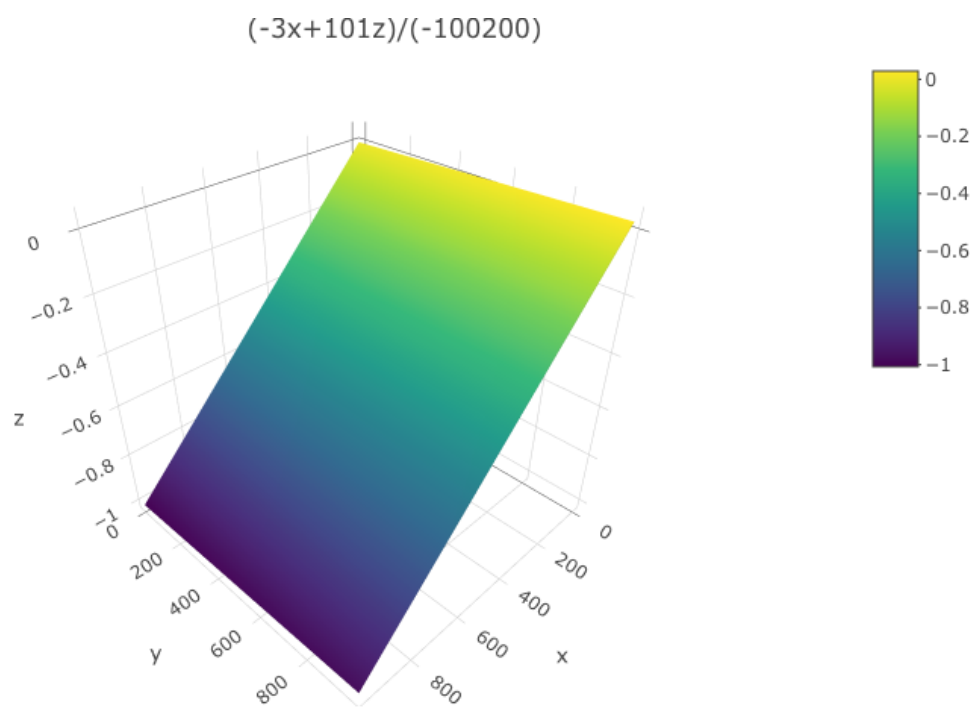


Figura R3

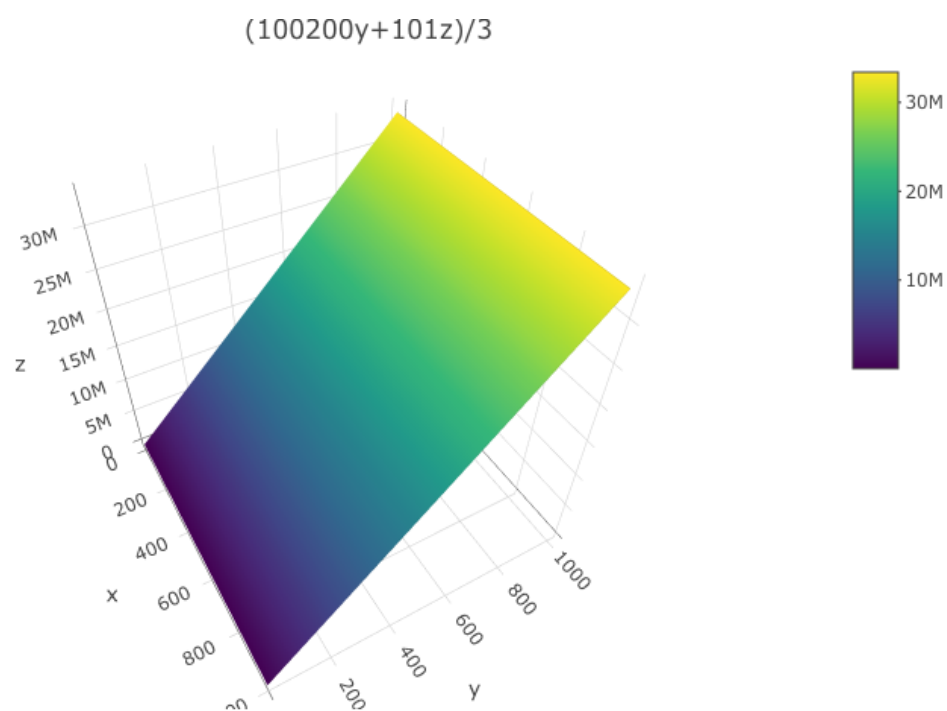
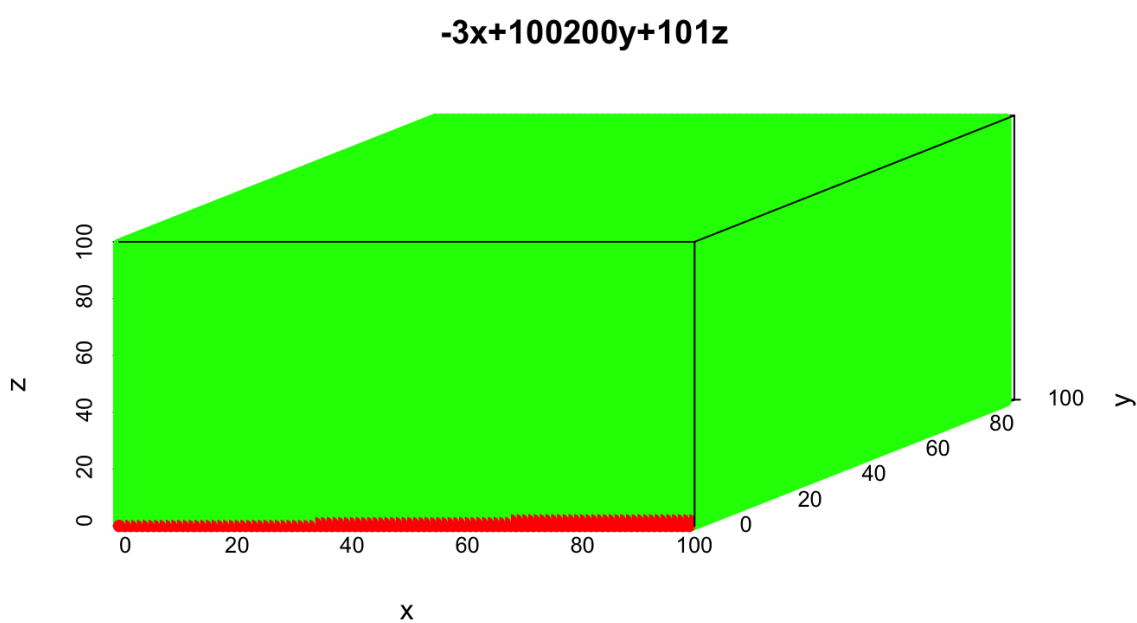


Figura R4



7.2. Analisi degli indici

Abbiamo eseguito l'analisi degli indici con le seguenti formule ottenute dai calcoli del costo degli indici $\log_2(x/6) - (\log_{512}(x/6) + 1) > 0$, per **data_visita_storico_indice** e **data_visita_corrente_indice**

nelle tabelle Visista_Veterinaria_Storico e Visista_Veterinaria_Corrente, e $\log_2(x/44) - (\log_{512}(x/44) + 1) > 0$, per **data_entrata_indice** e **data_uscita_indice** nella tabella Esempare_Animale, al fine di valutare meglio la convenienza di tali indici. La variabile x è il numero di righe presenti nella tabella considerata.

Abbiamo creato due grafici per le due formule:

- **I1**: risulta quasi sempre conveniente avere gli indici, l'unica parte in cui non conviene è per piccoli valori di X (sotto 1000)
- **I2**: risulta quasi sempre conveniente avere gli indici, l'unica parte in cui non conviene è per piccoli valori di X (sotto 1000)

Figura I1

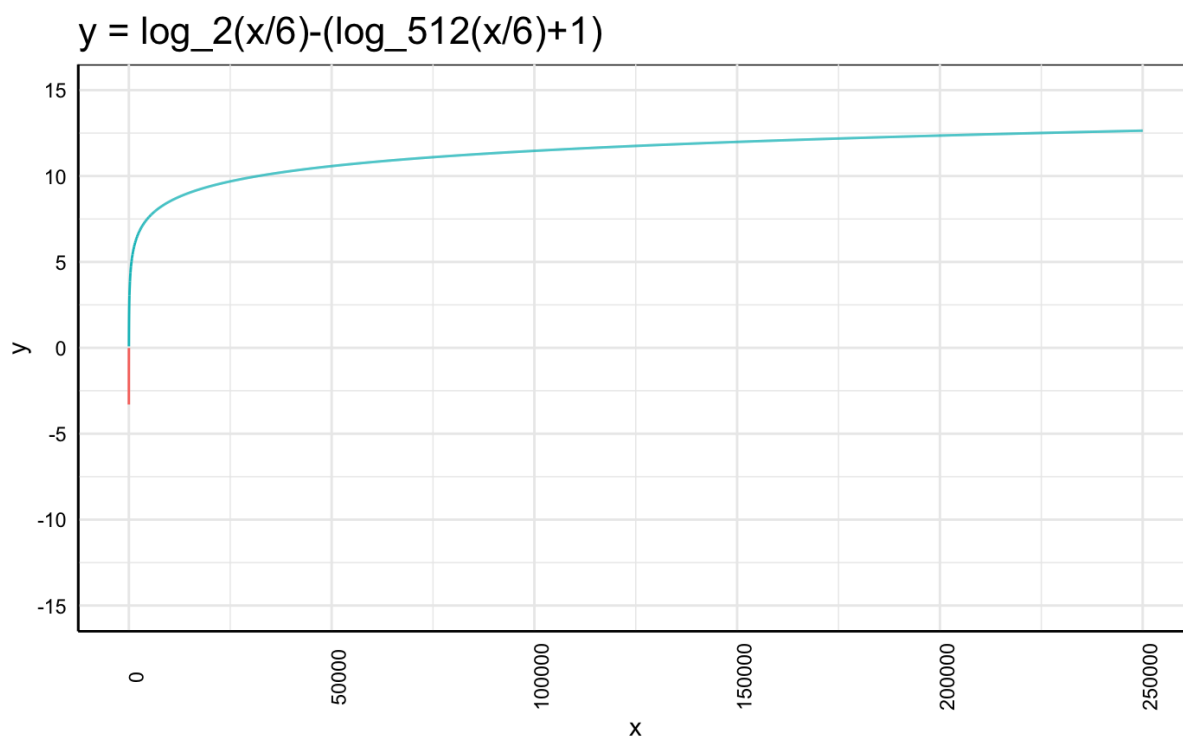
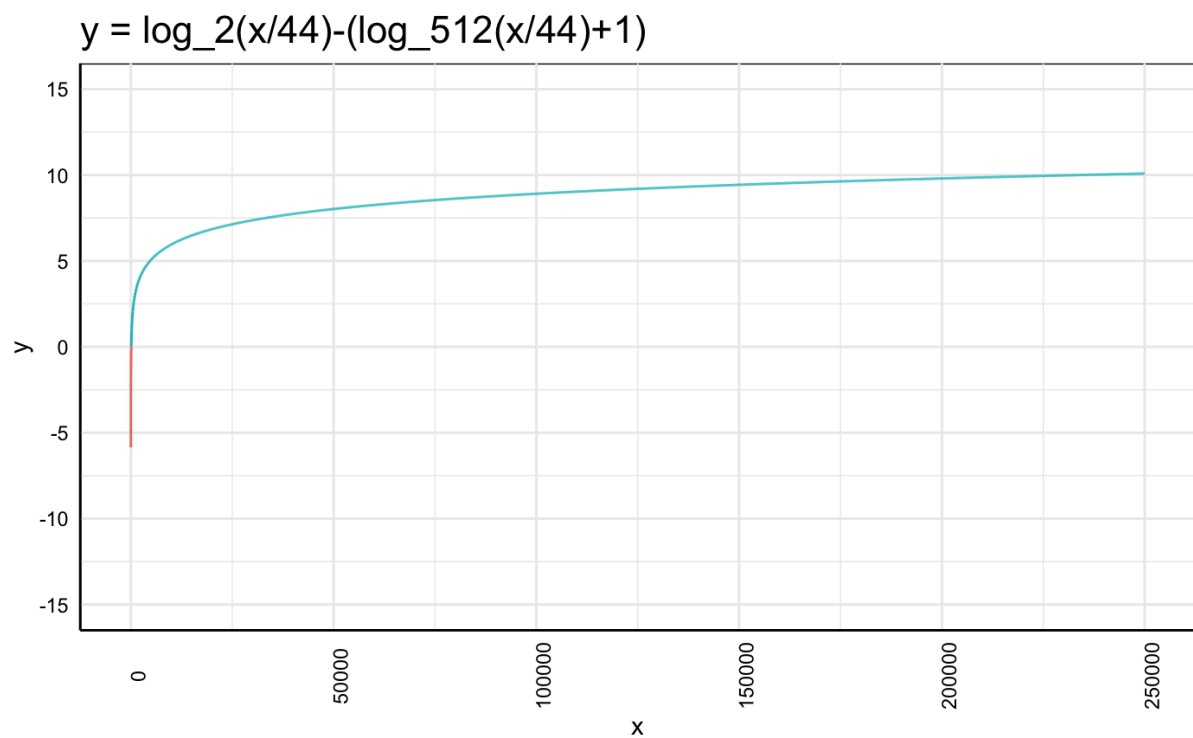


Figura I2



7.3. Esecuzione di alcune query significative

Abbiamo eseguito le query definite in precedenza (solo le query vere e proprie, non le insert o delete). Per le tabelle contenenti più di 30 righe, sono mostrate solo le prime 30 per ragioni di spazio.

Operazione 1

	data	animale	specie	malattia	peso	dieta	veterinario
1	2023-02-06	14	Snake unidentified		66409		VWGGON61Q76R110T

Operazione 3

	codice	paese_provenienza	Sesso	nome	data_nascita	data_entrata	data_uscita	specie	numero_visite_veterinarie	area
1	18	Indonesia	m	Abbey	2011-10-07	2021-02-18	NA	Greater sage grouse	25	Blue
2	7	Peru	f	Xena	2015-09-14	2021-01-02	NA	Kangaroo western grey	28	Blue
3	9	Colombia	f	Billy	1984-05-06	2021-01-25	NA	Kangaroo western grey	27	Blue
4	13	Russia	f	Robert	1979-05-25	2021-07-08	NA	Kangaroo western grey	22	Blue
5	15	Poland	m	Hazlett	1991-01-03	2021-02-22	NA	Kangaroo western grey	27	Blue
6	2	China	m	Reidar	1972-10-26	2021-11-13	NA	Kelp gull	18	Blue
7	4	Philippines	f	Gretna	1978-06-14	2021-12-15	NA	Kelp gull	18	Blue
8	6	Belarus	f	Shannon	2007-10-27	2021-07-03	NA	Kelp gull	21	Blue
9	13	New Zealand	f	Sigfried	1993-03-05	2021-01-29	NA	Kelp gull	27	Blue
10	5	China	f	Isabella	2014-03-03	2021-11-09	NA	Goldeneye common	19	Blue
11	7	Mexico	f	Cletus	2001-04-15	2021-07-15	NA	Goldeneye common	22	Blue
12	10	Tanzania	m	Nance	1987-12-15	2021-05-27	NA	Goldeneye common	24	Blue
13	19	United States	f	Isabella	1998-01-08	2021-01-17	NA	Goldeneye common	29	Blue
14	20	Colombia	f	Reube	1979-09-02	2021-08-13	NA	Goldeneye common	20	Blue
15	1	Nicaragua	f	Shamus	1979-12-08	2021-12-12	NA	Goose greylag	17	Blue
16	4	Chile	f	Minni	1989-01-23	2021-07-03	NA	Goose greylag	22	Blue
17	6	Philippines	m	Keenan	1984-08-27	2021-11-22	NA	Goose greylag	16	Blue
18	18	Portugal	f	Brandea	2016-10-12	2021-08-03	NA	Goose greylag	21	Blue
19	19	Zimbabwe	m	Gilli	2020-03-31	2021-10-08	NA	Goose greylag	19	Blue
20	2	China	f	Phebe	1992-07-18	2021-05-07	NA	Great egret	23	Blue
21	3	Rwanda	f	Cynde	1983-04-12	2021-04-09	NA	Great egret	24	Blue
22	4	China	m	Berna	1985-08-22	2021-06-10	NA	Great egret	23	Blue
23	6	China	m	Jeche	1980-07-20	2021-02-27	NA	Greater sage grouse	26	Blue
24	11	Brazil	m	Moise	1975-01-01	2021-11-16	NA	Greater sage grouse	16	Blue
25	6	Indonesia	m	Mignonne	2010-07-31	2021-12-15	NA	Pale white-eye	16	Green
26	13	Vietnam	m	Frederic	1980-06-10	2021-08-18	NA	Pale white-eye	20	Green
27	14	Czech Republic	f	Stephana	2020-09-26	2021-05-22	NA	Pale white-eye	23	Green
28	18	China	f	Brynne	2007-06-16	2021-05-02	NA	Pale white-eye	23	Green
29	1	China	f	Bronson	1983-09-26	2021-01-04	NA	Pine siskin	28	Green
30	5	Tunisia	m	Nissy	2011-09-08	2021-03-03	NA	Pine siskin	26	Green

Operazione 4

	gabbia	casa	area
	No data available in table		

Operazione 6

	gabbia	casa	area
1	d	1	Yellow
2	e	1	Yellow
3	d	2	Yellow
4	e	2	Yellow
5	e	3	Yellow
6	d	4	Yellow
7	e	4	Yellow
8	a	5	Yellow
9	b	5	Yellow
10	c	5	Yellow
11	d	5	Yellow
12	e	5	Yellow
13	d	6	Yellow
14	e	6	Yellow
15	e	7	Yellow
16	d	3	Blue
17	e	3	Blue
18	d	4	Blue
19	e	4	Blue
20	e	5	Blue
21	e	6	Blue
22	e	2	Green
23	c	3	Green
24	d	3	Green
25	e	3	Green
26	e	4	Green
27	e	5	Green
28	d	6	Green
29	e	6	Green
30	e	7	Green

Operazione 7

	↕	codice	↕	paese_provenienza	↕	Sesso	↕	nome	↕	data_nascita	↕	data_entrata	↕	data_uscita	↕	specie	↕	numero_visite_veterinarie	↕	data	↕
1		1		Mongolia		f		Hadlee		1979-12-23		2021-11-30		2022-05-04		Southern black-backed gull		6		2022-04-29	
2		1		Germany		m		Cece		2005-04-24		2021-07-21		2023-01-08		Rattlesnake eastern diamondback		19		2022-12-17	
3		1		China		m		Lewes		1991-04-05		2021-11-01		2023-01-06		Onager		16		2023-01-01	
4		1		Indonesia		f		Lemmie		1995-06-20		2021-09-18		2023-01-06		Arboreal spiny rat		17		2022-12-29	
5		1		Brazil		m		Manolo		2012-11-25		2021-08-09		2022-05-24		Greater sage grouse		12		2022-05-19	
6		1		Brazil		m		Letisha		1993-01-18		2021-05-23		2023-01-06		Coot red-knobbed		21		2022-12-17	
7		1		Brazil		m		Marjy		2001-08-21		2021-06-09		2023-01-06		Dragon omate rock		21		2022-12-22	
8		1		Argentina		m		Teresa		1978-08-27		2021-04-14		2022-04-27		American alligator		13		2022-04-09	
9		1		China		m		Law		1986-06-19		2021-04-29		2022-06-11		Brazilian tapir		14		2022-05-24	
10		1		Peru		m		Janka		2006-07-06		2021-12-01		2023-02-16		Kelp gull		18		2023-02-08	
11		1		Ukraine		m		Samson		1991-06-26		2021-10-31		2023-01-06		African buffalo		16		2022-12-30	
12		1		Burkina Faso		m		John		2006-07-25		2021-03-02		2023-01-06		Goldeneye common		24		2023-01-05	
13		1		Sweden		f		Stephenie		2000-10-08		2021-01-11		2022-11-11		Snowy owl		24		2022-11-07	
14		1		Poland		f		Nelli		1998-03-17		2021-01-06		2023-01-06		Screamer southern		23		2022-12-09	
15		1		Costa Rica		m		Flinn		2017-06-13		2021-07-13		2022-04-01		Constrictor eastern boa		11		2022-03-30	
16		1		United States		m		Stephi		2013-07-06		2021-11-30		2022-11-30		Andean goose		14		2022-11-17	
17		1		Cuba		m		Colver		2012-07-22		2021-01-09		2023-01-17		Radiated tortoise		25		2022-12-30	
18		1		Peru		m		Hendrick		1999-02-19		2021-12-13		2023-02-10		Cat long-tailed spotted		1		2021-12-13	
19		1		Bosnia and Herzegovina		m		Hubey		2001-01-24		2021-11-20		2022-07-13		Rattlesnake horned		10		2022-06-28	
20		1		Bosnia and Herzegovina		m		Domeniga		2016-04-19		2021-04-06		2022-09-12		Mockingbird galapagos		18		2022-08-29	
21		1		Portugal		f		Carlyn		2017-09-18		2021-01-19		2022-04-25		Buttermilk snake		16		2022-04-03	
22		1		Bulgaria		m		Lulu		1976-01-30		2021-10-01		2022-12-25		Red-cheeked cordon bleu		15		2022-11-25	
23		1		Indonesia		f		Ellary		2020-09-28		2021-03-09		2023-01-09		Bird pied butcher		24		2023-01-08	
24		1		Indonesia		m		Cosimo		2020-10-22		2021-07-19		2022-02-22		Anteater giant		8		2022-02-14	
25		1		Nicaragua		f		Shamus		1979-12-08		2021-12-12		NA		Goose greylag		17		2023-02-13	
26		1		China		f		Jeannine		1978-08-16		2021-07-17		2022-03-24		Kangaroo western grey		9		2022-02-26	
27		1		Egypt		f		Parnell		1981-09-07		2021-09-10		2022-09-26		Worm snake unidentified		14		2022-08-27	
28		1		China		f		Silvia		1998-03-20		2021-12-26		2022-04-10		Mynah indian		4		2022-03-26	
29		1		China		f		Bronson		1983-09-26		2021-01-04		NA		Pine siskin		28		2023-02-13	
30		2		Guatemala		m		Val		1981-09-22		2021-11-13		2022-04-25		White-fronted capuchin		6		2022-04-12	

Operazione 8

	↕	specie	↕	avg	↕
1		Andean goose		20.25000	
2		Kelp gull		17.65000	
3		Screamer southern		21.65000	
4		Flicker campo		22.80000	
5		Rattlesnake horned		20.31579	
6		Worm snake unidentified		19.10526	
7		Brazilian tapir		17.85000	
8		Constrictor eastern boa		21.55000	
9		Tenrec tailless		22.73684	
10		Red-cheeked cordon bleu		18.89474	
11		Arboreal spiny rat		21.10000	
12		Mynah indian		21.57895	
13		Bird pied butcher		21.52632	
14		Snake unidentified		20.45000	
15		Python carpet		20.00000	
16		Racer snake		16.94737	
17		Goose greylag		19.78947	
18		Goldeneye common		20.75000	
19		Great egret		18.60000	
20		American alligator		20.00000	
21		Rattlesnake eastern diamondback		20.15000	
22		White-fronted capuchin		19.80000	
23		Lizard desert spiny		19.80000	
24		Mockingbird galapagos		19.26316	
25		Glider feathertail		22.10526	
26		Roadrunner greater		18.60000	
27		Wapiti elk		22.31579	
28		Western bearded dragon		22.70000	
29		Buttermilk snake		19.00000	
30		Radiated tortoise		22.70000	

Operazione 9

	↕	codice	↕	paese_provenienza	↕	Sesso	↕	nome	↕	data_nascita	↕	data_entrata	↕	data_uscita	↕	specie	↕	? column?	↕
1		16		China		m		Garvin		1987-01-30		2021-08-04		2023-03-10		Badger american		1.019417	

Operazione 1 extra

	data	animale	specie	malattia	peso	dieta	veterinario
1	2022-11-26	9	Two-toed sloth		815679		DGNSAO74I34G132X
2	2021-12-01	9	Two-toed sloth		815673		RRSPAR95S04V869K
3	2022-01-30	9	Two-toed sloth		815675		OPFLBG35E28T343X
4	2021-10-02	9	Two-toed sloth		815674		UOQAQQ70P22F991L
5	2022-07-29	9	Two-toed sloth		815668		ZFOAKJ36I08K413L
6	2022-05-30	9	Two-toed sloth		815673		TXJSS76W27X542H
7	2022-03-31	9	Two-toed sloth		815672		OILQPX59S83T716L
8	2021-11-01	9	Two-toed sloth		815674		CBQNTJ86T75R285F
9	2022-08-28	9	Two-toed sloth		815680		QKPKQX43V87J697X
10	2022-09-27	9	Two-toed sloth		815670		GGGXXP31J75H693A
11	2022-06-29	9	Two-toed sloth	Peste suina africana	815668	Solo mangime medico per 4 settimane	PTLZUN37T49S915U
12	2021-12-31	9	Two-toed sloth		815677		RELZWK33P23V293V
13	2022-10-27	9	Two-toed sloth		815671		HUOOR58D37G603M
14	2022-03-01	9	Two-toed sloth		815672		FFNPLL70Q85Q417Y
15	2022-04-30	9	Two-toed sloth		815683		GTBBNJ43V27L642G
16	2022-12-26	9	Two-toed sloth		815676		HKASFO52W74K212M

Operazione 2 extra

	giorno_settimanale	codice_fiscale	nome	cognome	casa	area
1	Monday	NIFLXK68S67A926L	Derk	McAleese	1	Purple
2	Tuesday	NIJPEG72U00Q399T	Vera	MacVaugh	2	Purple
3	Wednesday	TPZPPF91M37K191E	Lil	Mougin	3	Purple
4	Thursday	APIUCT83R71G789I	Derrick	Sandeland	4	Purple
5	Friday	KACZUV72P71T726S	Ellis	Jiran	5	Purple
6	Saturday	XNULWLY87O65G340E	Lorene	Garrattley	6	Purple
7	Sunday	YSSHFM65X50V166O	Carmina	Groarty	7	Purple
8	Monday	XNULWLY87O65G340E	Lorene	Garrattley	1	Yellow
9	Tuesday	AGTUQZ92U34X028B	Caldwell	Whitear	2	Yellow
10	Wednesday	DUUTXH59L58S764F	Haley	Roseburgh	3	Yellow
11	Thursday	LLTSGX47P44G930D	Beverie	Ranshaw	4	Yellow
12	Friday	GXCEYQ91J34W302C	Sig	Gypson	5	Yellow
13	Saturday	NIFLXK68S67A926L	Derk	McAleese	6	Yellow
14	Sunday	NPSOTH18B53J305M	Carol	Durant	7	Yellow
15	Monday	GXCEYQ91J34W302C	Sig	Gypson	1	Blue
16	Tuesday	SLRGFS10D03P120Q	Dot	Battaille	2	Blue
17	Wednesday	XNULWLY87O65G340E	Lorene	Garrattley	3	Blue
18	Thursday	DUUTXH59L58S764F	Haley	Roseburgh	4	Blue
19	Friday	NIJPEG72U00Q399T	Vera	MacVaugh	5	Blue
20	Saturday	TPZPPF91M37K191E	Lil	Mougin	6	Blue
21	Monday	NGNEJS20G44O218U	Lem	Pisculli	1	Green
22	Tuesday	NPSOTH18B53J305M	Carol	Durant	2	Green
23	Wednesday	EQIDRV14A31P767M	Guntar	Wotherspoon	3	Green
24	Thursday	SLRGFS10D03P120Q	Dot	Battaille	4	Green
25	Friday	MOLJQW25F69I447J	Benjamin	Baty	5	Green
26	Saturday	NRSQFY91P40Y861R	Selia	Jutson	6	Green
27	Sunday	MOLJQW25F69I447J	Benjamin	Baty	7	Green
28	Monday	FAQKQA16W43S271M	Ophelia	Jellico	1	Red
29	Tuesday	BERTGA20V52K690G	Rivkah	Chatterton	2	Red
30	Wednesday	NIFLXK68S67A926L	Derk	McAleese	3	Red

7.4. Analisi dei dati presenti nella base di dati e altre analisi statistiche

Abbiamo eseguito l'analisi dei dati presenti nella base di dati, essendo generati in maniera casuale non riflettono una situazione reale.

Le analisi sono:

- **D1, D2 e D3:** gli esemplari di animale totali, presenti e usciti dallo zoo; in questo caso si può notare una distribuzione piuttosto omogenea, con valori reali è possibile riscontrare una distribuzione più variegata, **per esempio** specie di animali con longevità maggiore rispetto a specie di animali con longevità minore, nel corso delle generazioni saranno di meno, in quanto ci saranno più ricambi generazionali della specie con aspettativa di vita minore nello stesso periodo di tempo rispetto alle specie con aspettativa di vita maggiore
- **D4, D5 e D6:** le gabbie totali, occupate e vuote delle varie case; in questo caso si può notare una distribuzione variegata a causa del vincolo che ogni casa è dedicata ad una specie animale
- **D7, D8 e D9:** il numero di visite veterinarie di ogni esemplare di animale, media di ogni specie di animale e esemplare di animale rispetto alla media della specie in percentuale; in questo caso i valori sono piuttosto omogenei, con valori reali è possibile riscontrare una distribuzione più variegata, si potrebbe aggiungere un filtro ai grafici D7 e D9 mostrando gli esemplari di animale appartenenti ad una sola specie invece di visualizzare tutti gli esemplari di animale, **per esempio** può essere utile per identificare un esemplare di animale che ha eseguito un numero di visite veterinarie molto superiore alla media della sua specie di animale
- **D10:** rappresenta l'andamento del peso e le visite veterinarie eseguite a causa di malattie di un esemplare di animale; in un contesto reale questa analisi può servire per verificare una correlazione tra peso e malattia

Figura D1

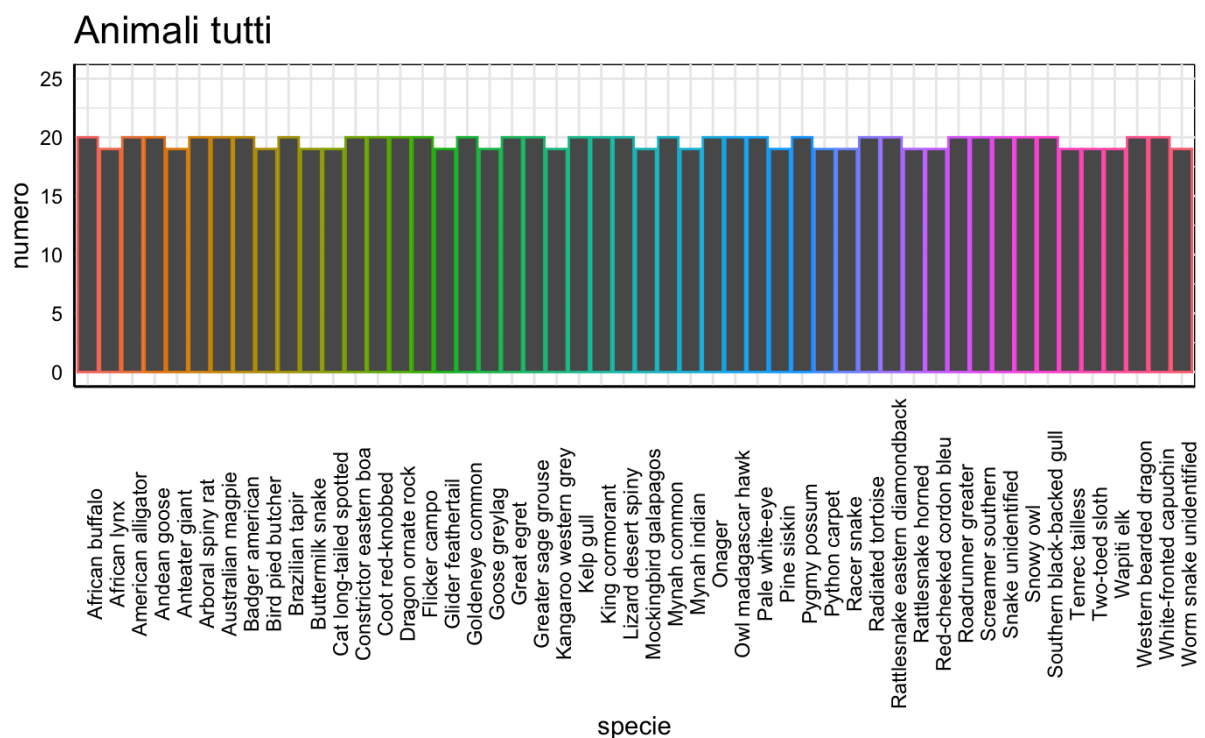


Figura D2

Animali presenti

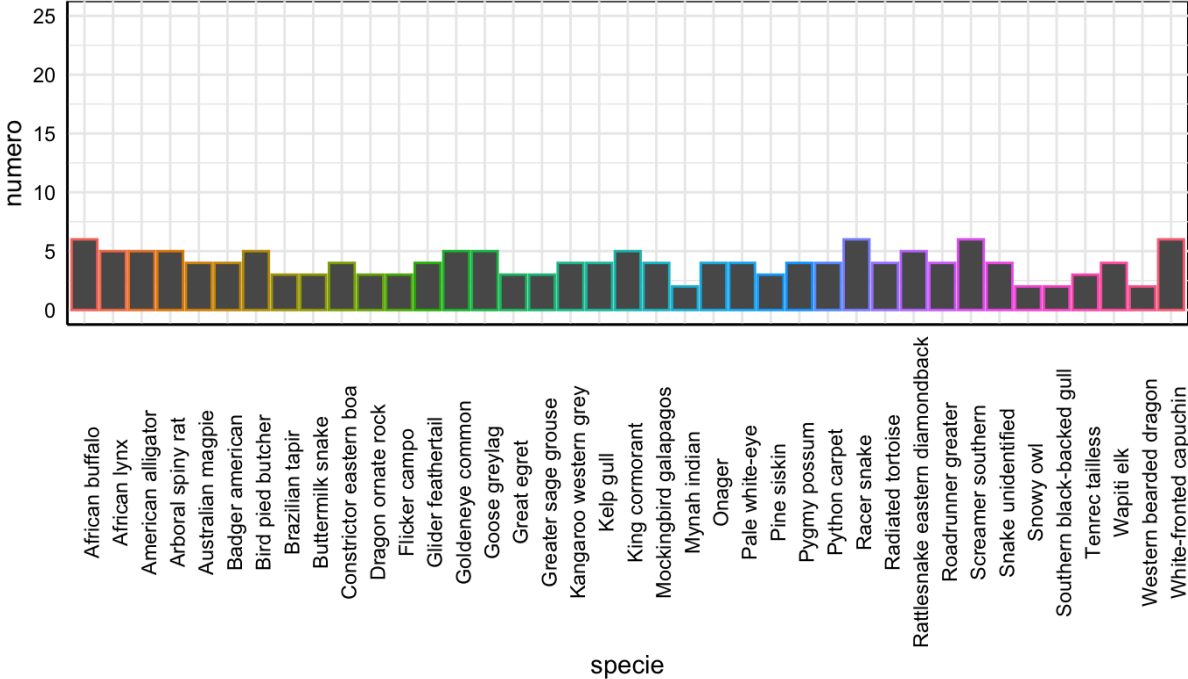


Figura D3

Animali usciti

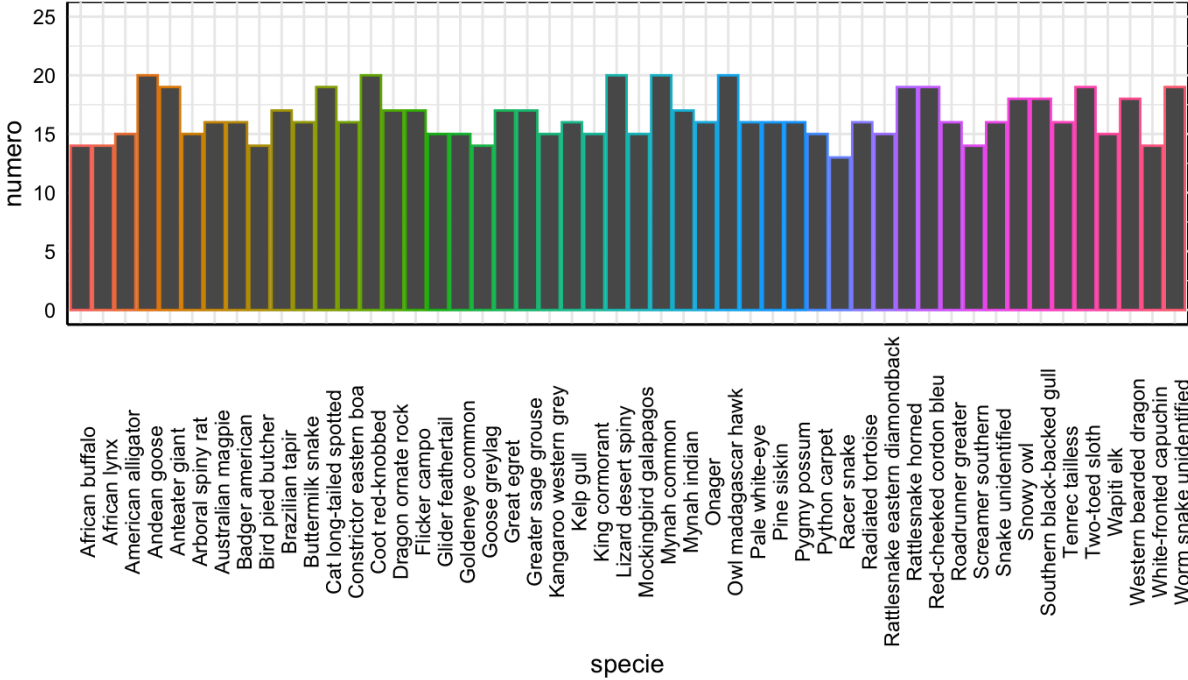


Figura D4

Gabbie totali

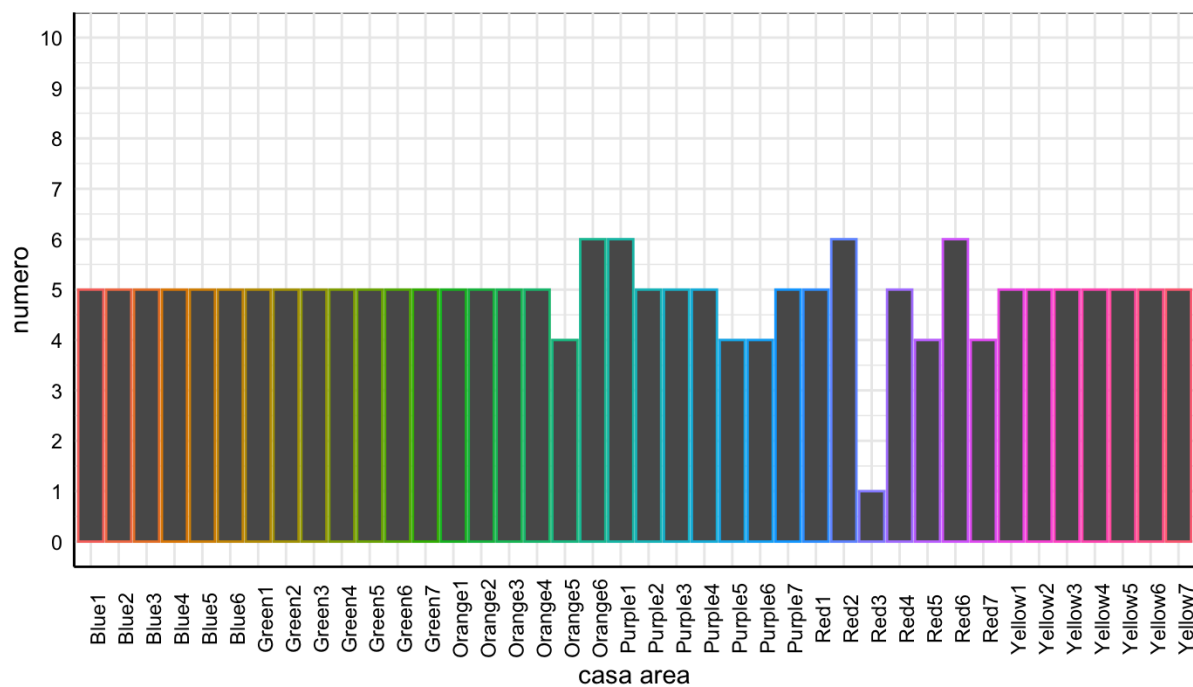


Figura D5

Gabbie occupate

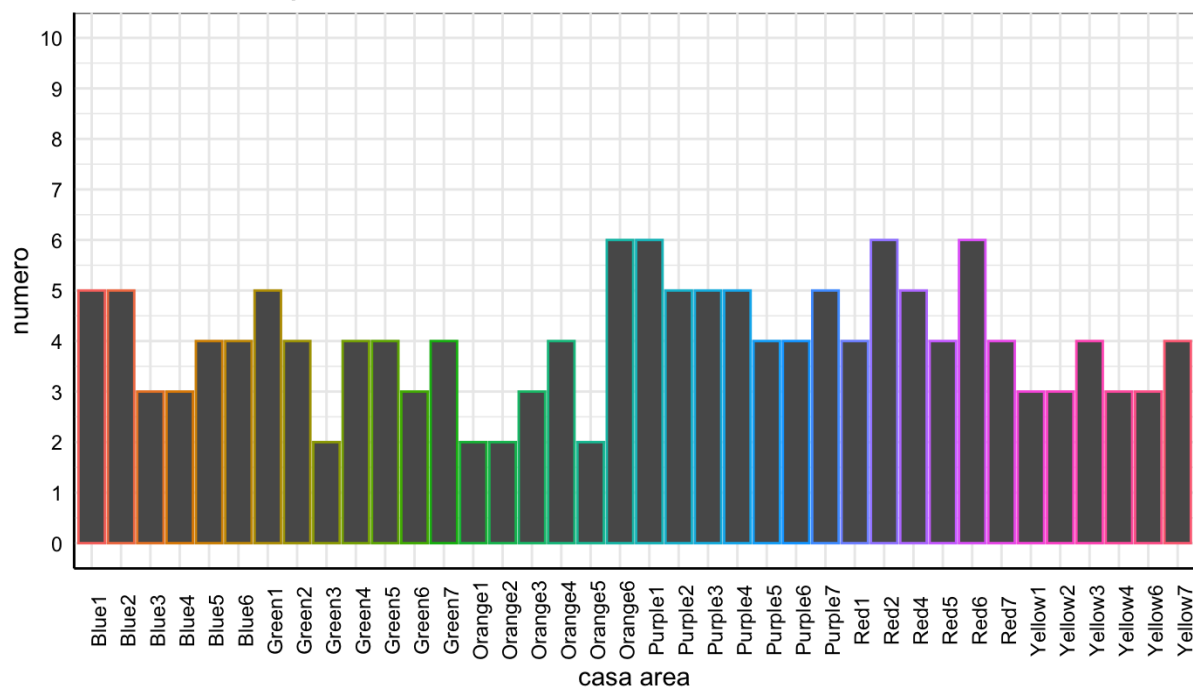


Figura D6

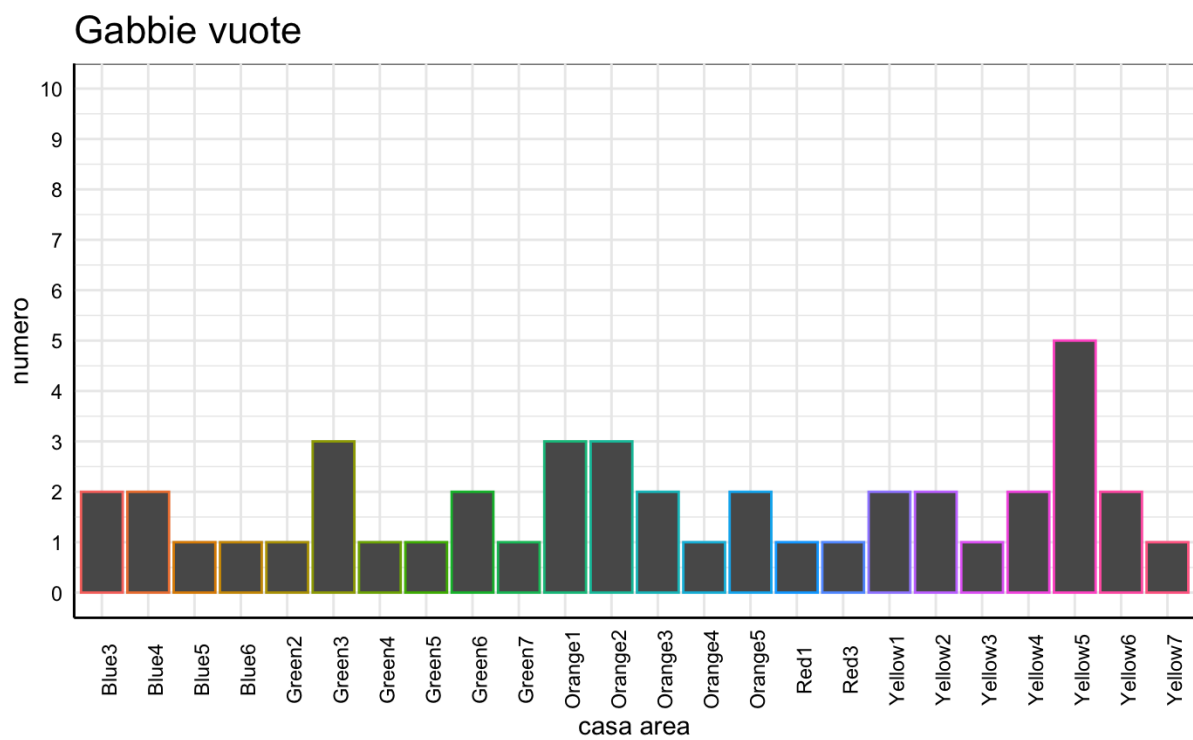


Figura D7

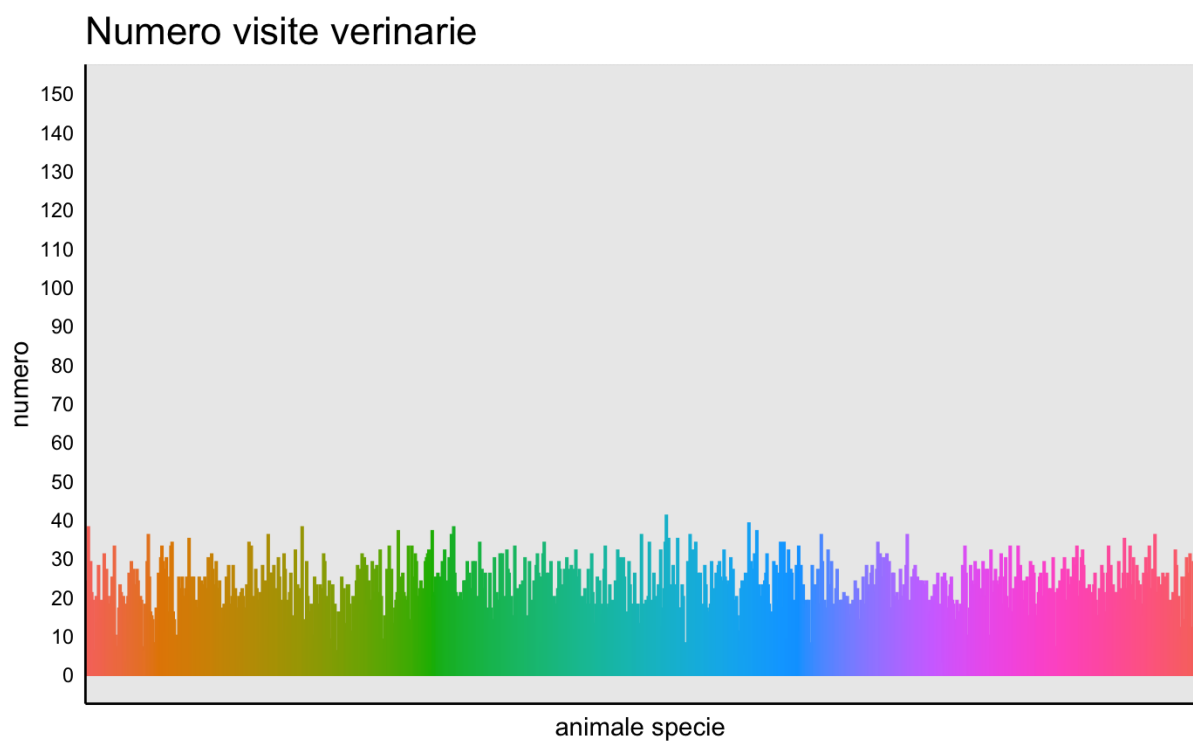


Figura D8

Numero visite verinarie - media specie

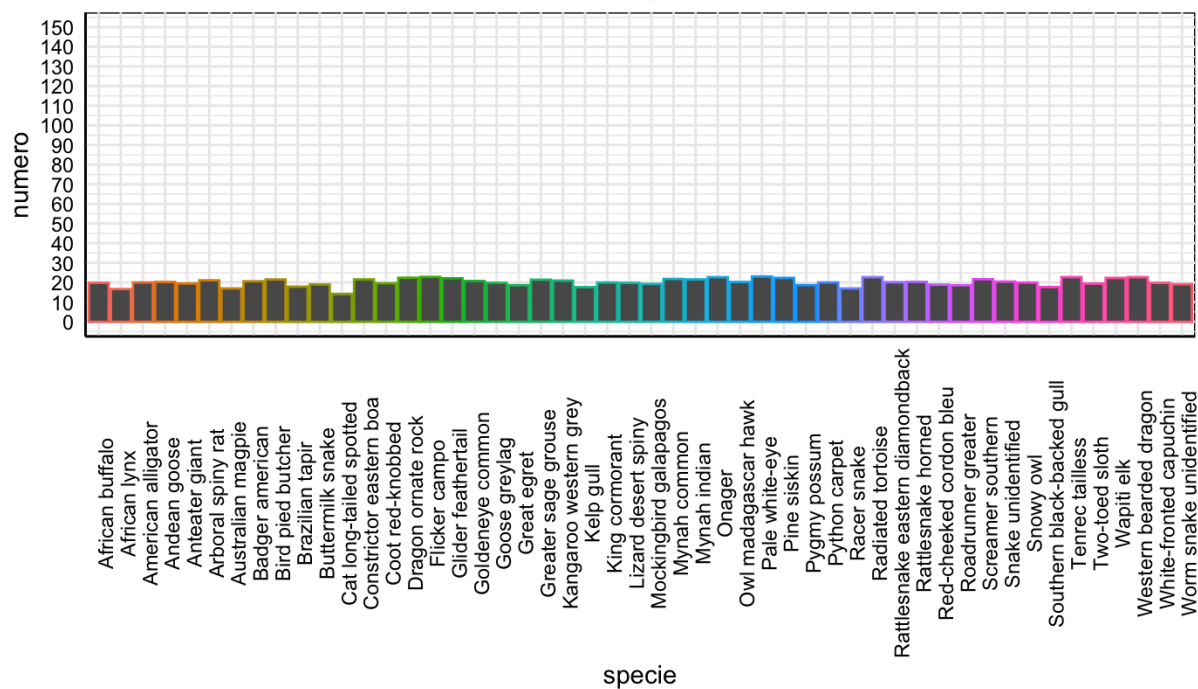


Figura D9

Numero visite verinarie - animale rispetto specie

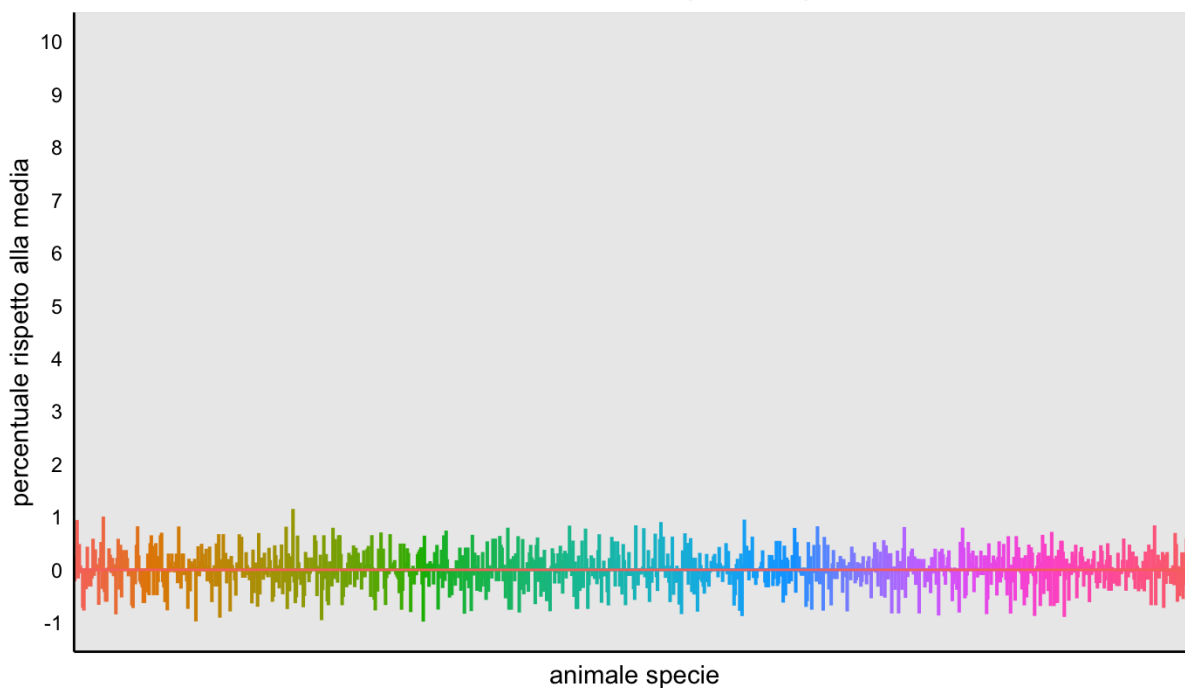


Figura D10

Salute di Keane

