

Article

An In-Depth Presentation of the ‘rhoneycomb’ R Package to Construct and Analyze Field-Experimentation ‘Honeycomb Selection Designs’

Anastasios Katsileros ¹, Nikolaos Antonetsis ¹, Maria-Georgia Gkika ¹, Eleni Tani ¹ , Penelope J. Bebeli ^{1,*} and Ioannis Tokatlidis ² 

¹ Laboratory of Plant Breeding & Biometry, Department of Crop Science, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece; katsileros@aua.gr (A.K.); stud610027@aua.gr (N.A.); stud117024@aua.gr (M.-G.G.); etani@aua.gr (E.T.)

² Department of Molecular Biology & Genetics, Democritus University of Thrace, Dragana Campus, 68100 Alexandroupolis, Greece; itokatl@mbg.duth.gr

* Correspondence: bebeli@aua.gr

Abstract: The Honeycomb Selection Design (HSD) is an innovative experimental method whose main feature is the even and systematic entry arrangement. Its systematicity, if combined with the absence of inter-plant competition that maximizes the phenotypic expression and differentiation of individual plants, enables the implementation of single-plant selection as early as the initial generations of genetic segregation, facilitating plant breeders to identify superior genotypes. Due to the specificity of entry allocation and the complexity of statistical data analysis, a specialized software becomes necessary. This article provides a detailed presentation of the ‘rhoneycomb’, a free and open-source R package concerning the construction, visualization, and analysis of HSDs.

Keywords: coefficient of response to selection; entry arrangement; homeostasis index; moving ring (circle or complete circular block); plant yield efficiency; predictive equation; spatial heterogeneity



Citation: Katsileros, A.; Antonetsis, N.; Gkika, M.-G.; Tani, E.; Bebeli, P.J.; Tokatlidis, I. An In-Depth Presentation of the ‘rhoneycomb’ R Package to Construct and Analyze Field-Experimentation ‘Honeycomb Selection Designs’. *Agronomy* **2023**, *13*, 2145. <https://doi.org/10.3390/agronomy13082145>

Academic Editors: George Menexes and Angelos Markos

Received: 24 July 2023

Revised: 12 August 2023

Accepted: 14 August 2023

Published: 16 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The purpose of a plant breeder is to develop and improve crop varieties for specific traits and characteristics. Breeders use various techniques to achieve their goals, encompassing traditional breeding methods and modern biotechnological tools [1]. In addition to the breeding methods and techniques, the field-experimentation model also plays an important role in the efficiency of genotype selection [2,3]. The appropriate field-experimentation model provides valuable insights into genetic variation, facilitates the identification of desirable traits, helps understand gene–environment interactions, and validates new breeding methods [4,5]. Plant breeders have several experimental methods at their disposal, and the choice of the appropriate method is critical at all stages of the breeding process [6].

An innovative field-experimentation model in crop breeding is the “Honeycomb Selection Design” (HSD), proposed by Prof. Fasoulas for single-plant selection in the absence of competition (i.e., the nil-competition regime) [7–10]. The main feature of HSDs is that individual plants are placed either in the center or vertices of a honeycomb pattern, and the inter-plant distance excludes plant-to-plant interference for any input. The growth of plants at nil-competition allows their maximum phenotypic expression and differentiation, which according to the breeder’s equation of expected response to selection [11], is a prerequisite for the successful selection of high-yielding and stable genotypes [12–14]. Furthermore, with a standardized, even, and systematic entry layout instead of the randomized configuration to implement the main principles met in other models, such as blocking, replication, and nearest neighbor adjustment on the same baseline, the HSD model is more efficient than the popular ones in reducing the experimental error [15]. Therefore, HSDs mitigate

the confounding effect of spatial heterogeneity on individual plant yields, i.e., the acquired part of the plant-to-plant variability owed to soil heterogeneity is inflated by uneven seed emergence, effects of clouds and capping in wet soils, uneven application of applied inputs, differential effects of herbivores, parasites or pathogens and interactions among these factors [13,15]. For all these reasons, HSDs are especially useful for evaluating heterogeneous progeny lines (entries) and applying single-plant selection within the superior ones, enabling breeders to identify genotypes that enhance crop productivity and stability potentially [15–17].

Although the advantages of the ‘Fasoulas’ method contribute to faster genetic progress, more reliable selection decisions, and the development of superior crop varieties with desirable traits [18], HSDs have certain drawbacks for incorporating them in crop breeding programs, such as requirements, limitations, and complexity. Due to enlarged inter-plant distance, HSDs require large areas, while it is not always easy to mechanize experimental establishment and harvesting. For the construction and statistical analysis of HSDs, two statistical software are available, i.e., the ‘Prognostic Breeding Application’ (JMP Add-In program) [19] and the ‘rhoneycomb’ package in R [20].

A detailed description of the HSDs is given by Fasoulas and Fasoula [21]. This paper aims to thoroughly present the ‘rhoneycomb’ package, a free and open-source statistical tool for constructing, visualizing, and analyzing HSDs. The first version of this package was introduced at the 18th Conference of the ‘Hellenic Scientific Society of Genetics and Plant Breeding’ [22]. We present a summary of HSDs’ general characteristics, the construction and analysis adapted to the ‘rhoneycomb’ R package, and then instructions for using this software. The upgraded version described here contains additional functions and new features in visualization, including empirical data utilized to illustrate the essential steps, thus exemplifying the package.

2. Overview of the Honeycomb Selection Designs

Advanced experimental HSDs enhance the efficiency of selecting high-yielding and stable genotypes, provided that individual plants grow without competition. The main feature of HSDs that plays a decisive role in the above achievement is the systematic rather than random entry arrangement to efficiently sample the spatial heterogeneity. They are distinguished into two categories, depending on whether the designs evaluate one or more entries.

Designs constructed for a single entry, represented by many plants, are denoted as HSD0. Optionally, they may include one (HSD01) or three entry-controls (HSD03), evenly allocated across the honeycomb array, to promote sampling of spatial heterogeneity in the comparison of plants, thus improving the reliability of the single-plant selection. The breeder can evaluate any plant (genotype) as potentially selectable in these designs.

On multiple entries, evenly and systematically arranged, designs are denoted as HSD(E), e.g., HSD(7) and HSD(21) for 7 and 21 entries, respectively. The breeder determines the number of plants per entry (replications) according to the nature of the entries and the expected heterogeneity (i.e., genetically heterogeneous entries and/or high spatial heterogeneity require an increased number of replications). The number of entries (E) included in a particular design is given by:

$$E = X^2 + XY + Y^2 \quad (1)$$

where, X and Y are called coordinates of the design and are integer numbers ($X + Y \geq 0$).

2.1. Construction of a Honeycomb Selection Design

Individual plants in any HSD are equally placed. The inter-plant distance (d), determined by the breeder, should be large enough to preclude plant-to-plant competition. Each plant occupies the center of an inner hexagonal experimental plot and is adjacent to six plants occupying the vertices of the hexagonal pattern. The experimental plot has sides of $d\sqrt{3}/3$ in length and corresponds to an area of $d^2\sqrt{3}/2$. The exact position of a plant

(entry) in the hexagonal experimental plot depends on the parameters and features of the specific HSD.

In a single-entry design (HSD0), plants are randomly allocated to the entire experimental area. Such a design may include either one (Figure 1a) or three (Figure 1b) entry-controls which are placed on specific plots and occupy one per seven experimental plots. The plots of an entry-control are determined by the K parameter (1, 3, 5 or 7). Different K values generate different control positions, however, neighboring positions of the same control always form an equilateral triangle, while their coordinates (X and Y) are equal to 2 and 1.

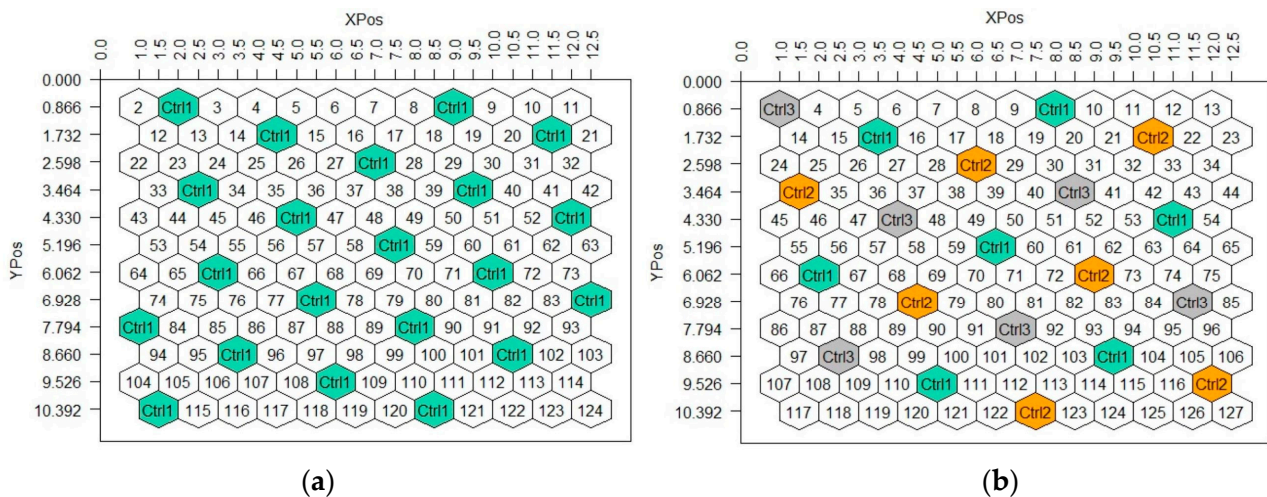


Figure 1. (a) A single-entry design with one entry-control ($X = 2$, $Y = 1$ and $K = 1$), and (b) a single-entry design with three entry-controls ($X = 2$, $Y = 1$ and $K = 1$) ($d = 1$ m).

The multiple-entry designs, HSD(E), can be classified into two types, ungrouped and grouped. Ungrouped designs feature the entire set of entries in each row, arranged sequentially (Figure 2a). Grouped designs distribute the entries in groups across multiple rows (Figure 2b). The distinction between a grouped and ungrouped design can be determined by examining the numerical values of its X and Y coordinates. If the X and Y coordinates of the design do not share any common factors (divisors) other than 1, then the design is classified as ungrouped. Conversely, if a common factor (divisor) exists other than 1 between the X and Y coordinates, the design is classified as grouped. The common factor of the grouped designs determines the number of groups to which E is divided, whereas the ratio $E/\text{divisor}$ determines the size m of groups. Furthermore, in both types of HSD(E) the coordinates X and Y indicate the distance in plant-to-plant steps (X -horizontal and Y -diagonal and vice versa) between the two nearest neighboring plants that belong to the same entry (Figure 2).

To construct an HSD(E), it is necessary to determine the initial entry code (number) for each field row. The initial number of each row is unique, specified by the parameter K_n ($n = 1, 2, \dots, 6$). The algorithms for calculating the parameters K_n and the initial number for each row are presented by Fasoulas and Fasoula [21]. Ungrouped designs have two constants (K_1 and K_2), whereas grouped designs have four constants (K_3 , K_4 , K_5 and K_6). For a particular HSD(E), the breeder chooses any of the corresponding K_n parameters. Although different K_n parameters generate different initial numbers resulting in different entry allocation, the principles of even and systematic entry arrangement are not disturbed.

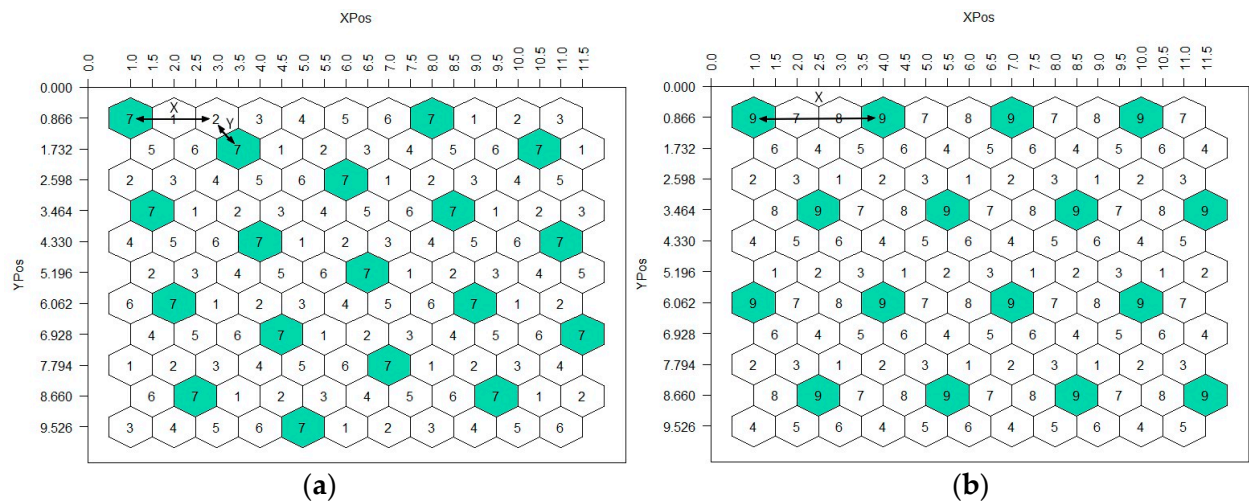


Figure 2. (a) The ungrouped HSD(7) with seven entries (E7, $X = 2$, $Y = 1$ and $K = 2$), and (b) the grouped HSD(9) with nine entries (E9, $X = 3$, $Y = 0$, $K = 2$ and $m = 3$); in both designs $d = 1$ m (product of the ‘rhoneycomb’ package according to Fasoulas and Fasoula [21]).

Generally, parameters K are used for the construction of HSDs. Any K of those available for a particular HSD arranges the entries evenly and systematically across the experimental area. In constructing of a particular HSD, different K values obviously result in different starting numbers of the rows. Still, in all resulting arrays, entries with similar numbers will form the characteristic equilateral triangle pattern (Figure 3b), and each entry will occur in the center of a moving, quasi-circular, and complete replicate [21]. In other words, different K values generate different entry positions, however, any entry is constantly surrounded by all the rest entries, i.e., occupies the center of a circular skeleton of a certain size that contains the entire set of entries, thus forming a complete circular block. Therefore, it does not matter which K the breeder will use to construct a particular HSD.

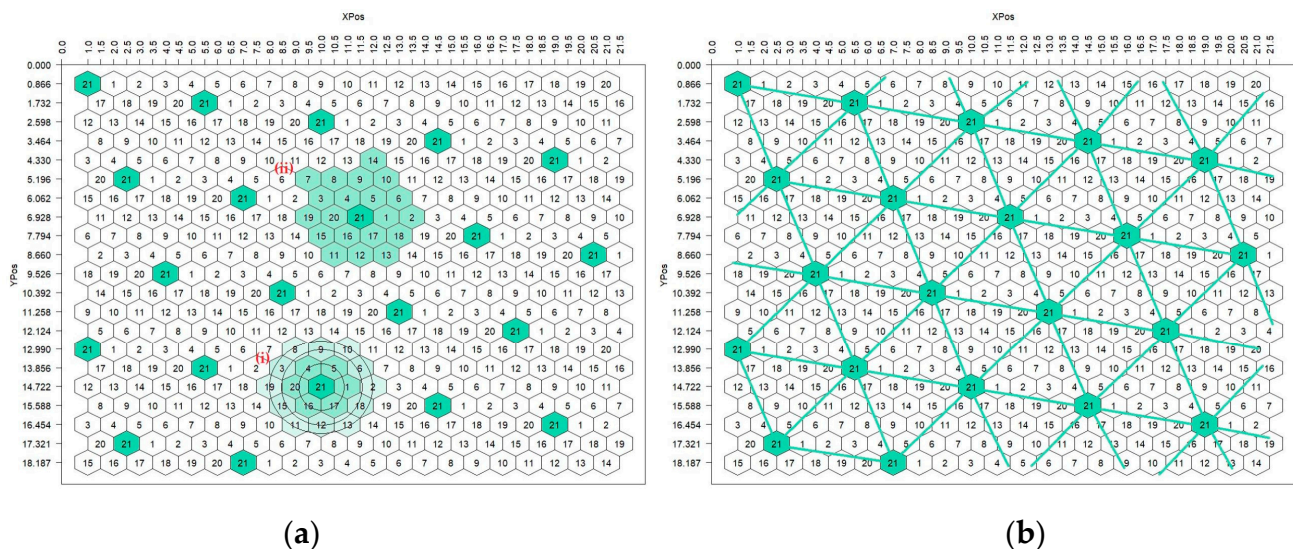


Figure 3. The ungrouped HSD(21) ($K = 4$) demonstrates (a) the moving ring as either a (i) circle (size = 6, 12, 18 or more plants on the circumference) or (ii) a circular skeleton containing all the entries (complete circular block), and (b) the triangular grid for entry 21.

2.2. Analysis of a Honeycomb Selection Design

The HSD's statistical analysis relies on the utilization of the circular moving ring. There are two types of moving ring. The commonly used one is a circle of flexible size, formed by the nearest neighboring plants according to their radius-distance from the central plant (Figure 3(ai)). The circle size is determined by the number of circumference plants which can vary from 6 to more than 120 plants. The breeder chooses a particular circle size, depending on the nature of entries and the experimental conditions. A larger circle is chosen when entries are highly heterogeneous (e.g., segregating progeny sister lines) and when there is high spatial heterogeneity or a high rate of missing plants in the experiment. The alternative type of moving ring is a circular skeleton of a certain size, since it contains the entire set of entries (Figure 3(aii)), thus forming a complete circular block. Due to the triangular pattern of entry allocation (Figure 3b), the moving ring, having a specific plant in the center, consistently covers the entire experimental area, thus constituting a block replication. This approach samples the spatial heterogeneity in the most efficient way to mitigate its obscuring effects. Hence, breeders can identify the outstanding entries and select the superior individual plants (genotypes) using the relevant indices and predictive equations [23].

The main selection criterion is the genotype's ability to respond to resources. It is reflected by the grain yield per plant at the nil-competition regime, defined as Plant Yield Efficiency (PYE) [23]. The respective statistical measure, named PYE index, leverages the principle of the nearest neighbor adjustment corresponding to the squared relative yield of the central plant of the moving ring and is calculated by the equation:

$$PYE = (x/\bar{x}_r)^2 \quad (2)$$

where, x is the absolute yield of the central plant and \bar{x}_r is the mean yield of the moving ring. Individual plants are ranked for their PYE values, while the average entry PYE (mPYE) is used to distinguish the most promising entries, and within high mPYE entries individual high PYE plants are selected.

Another selection criterion is the genotype's ability to withstand the environmentally induced acquired plant-to-plant variation as a stability measure. It is approached by the entry Homeostasis Index (HI):

$$HI = (\bar{x}/s)^2 \quad (3)$$

where \bar{x} and s denote the entry's mean yield and standard deviation, respectively.

To apply single-plant selection, the breeder should focus on entries of high values in both mPYE and HI. However, it is recommended the HI is cautiously considered in early segregating generations due to genetic heterogeneity. The HI may undervalue a highly heterogeneous progeny sister line that reasonably would have a low HI value [24]. The breeder should focus exclusively on mPYE in highly heterogeneous entries.

Considering a plant for selection, its PYE and HI of the entry it belongs can be combined into a unique equation, the Plant Predictive Equation (PPE):

$$PPE = PYE \cdot HI = (x/\bar{x}_r)^2 \cdot (\bar{x}/s)^2 \quad (4)$$

If the breeder wishes to combine the two indexes at the entry level, the mean Plant Predictive Equation (mPPE) is calculated, i.e., the average entry PPE value. In another option, the first part of Equation (4) is substituted by the Entry Yield Index (EYI) which is the squared relative entry mean yield $(\bar{x}/\bar{x}_t)^2$, where \bar{x}_t is the grand mean of the trial, resulting in the Entry Predictive Equation (EPE):

$$EPE = EYI \cdot HI = (\bar{x}/\bar{x}_t)^2 \cdot (\bar{x}/s)^2 \quad (5)$$

Entries are ranked in descending order of either mPPE or EPE to identify superior ones within which selection can be applied.

To establish a breeding program, the breeder may wish to evaluate several available germplasms, e.g., populations or landraces, to focus on the most promising ones as starting material [25,26]. For this particular purpose, germplasms are tested together in an HSD, including a homogeneous entry-control (e.g., an elite cultivar), having a special statistic tool developed for the first time to predict the response to single-plant selection, the Coefficient of Response to Selection (CRS):

$$CRS = (\bar{x}_{sel} - \bar{x}_{contr}) / s \quad (6)$$

where \bar{x}_{sel} is the average yield of a predetermined number of selectable plants (those of the highest PYE) of the germplasm under consideration and s is its standard deviation, while \bar{x}_{contr} is the mean yield of an entry-control. The higher the CRS value of germplasm, the greater its potential to include superior genotypes.

3. Usage of the ‘rhoneycomb’ Package

The ‘rhoneycomb’ package comprises a set of functions intended to aid plant breeders in constructing and visualizing honeycomb selection designs, conducting statistical analysis on experimental data, and presenting the results. For developing the ‘rhoneycomb’ package, several open and free software, such as R and packages devtools, stats, utils and graphics have been used [27,28]. The package has been tested under Windows, Linux, and MacOS environments and is published in the Comprehensive R Archive Network (CRAN) at the link: <https://cran.r-project.org/web/packages/rhoneycomb/index.html> (accessed on 26 June 2023).

3.1. Installation of the Package

To install the ‘rhoneycomb’ package in R, the user executes the provided command within the R environment console:

```
> install.packages("rhoneycomb")
```

Alternatively, if the user uses the RStudio desktop, a friendly and comprehensive development environment for R [29], the package can be installed with the installer (Figure 4).

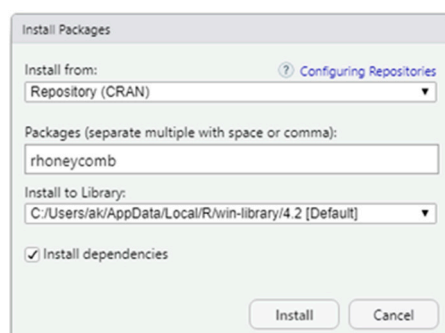


Figure 4. Installing the ‘rhoneycomb’ package using the installer of Rstudio.

To start utilizing the package after installation, the user must load it into an R session by executing the following command.

```
> library(rhoneycomb)
```

The package automatically checks for the required R packages (dependencies).

3.2. Generate Available Designs

By implementing the algorithmic steps outlined in Fasoulas and Fasoula [21], the function generate() was developed. This function generates the essential parameters and features required to construct the HSD. Within the function, the E_gen argument is available to use. Users enter in the E_gen argument a single number or a vector of entries. To generate

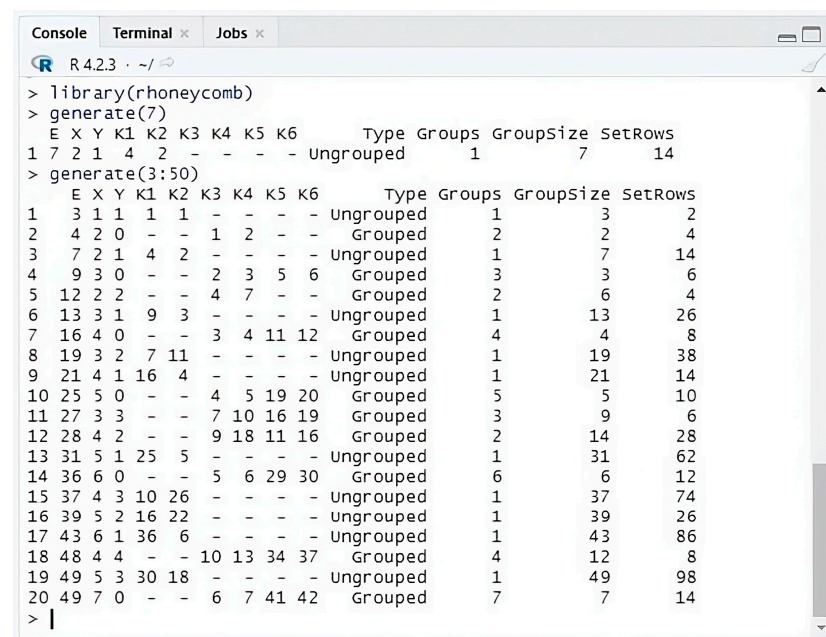
the parameters and features, for example for the HSD(7), the user executes the following command:

```
> generate(7)
```

Alternatively, the user executes a command including a range of entries (e.g., 3 to 50 entries) to reveal the available designs:

```
> generate(3:50)
```

These functions return a data frame with the available HSDs (number of entries E), the coordinates (X and Y) and K_n parameters of the design, the type of the design (ungrouped or grouped), the number of groups (Groups), the size of the groups (GroupSize), and the number of rows (SetRows) required for one complete iteration of the design (Figure 5).



	E	X	Y	K1	K2	K3	K4	K5	K6	Type	Groups	GroupSize	SetRows
1	7	2	1	4	2	-	-	-	-	Ungrouped	1	7	14
2	3	1	1	1	1	-	-	-	-	Ungrouped	1	3	2
3	4	2	0	-	-	1	2	-	-	Grouped	2	2	4
4	7	2	1	4	2	-	-	-	-	Ungrouped	1	7	14
5	9	3	0	-	-	2	3	5	6	Grouped	3	3	6
6	12	2	2	-	-	4	7	-	-	Grouped	2	6	4
7	13	3	1	9	3	-	-	-	-	Ungrouped	1	13	26
8	16	4	0	-	-	3	4	11	12	Grouped	4	4	8
9	19	3	2	7	11	-	-	-	-	Ungrouped	1	19	38
10	21	4	1	16	4	-	-	-	-	Ungrouped	1	21	14
11	25	5	0	-	-	4	5	19	20	Grouped	5	5	10
12	27	3	3	-	-	7	10	16	19	Grouped	3	9	6
13	28	4	2	-	-	9	18	11	16	Grouped	2	14	28
14	31	5	1	25	5	-	-	-	-	Ungrouped	1	31	62
15	36	6	0	-	-	5	6	29	30	Grouped	6	6	12
16	37	4	3	10	26	-	-	-	-	Ungrouped	1	37	74
17	39	5	2	16	22	-	-	-	-	Ungrouped	1	39	26
18	43	6	1	36	6	-	-	-	-	Ungrouped	1	43	86
19	48	4	4	-	-	10	13	34	37	Grouped	4	12	8
20	49	5	3	30	18	-	-	-	-	Ungrouped	1	49	98
21	49	7	0	-	-	6	7	41	42	Grouped	7	7	14

Figure 5. The parameters and features of several HSDs.

3.3. Construction and Visualization of Honeycomb Selection Designs

Four functions are available to facilitate the construction and visualization of honeycomb selection designs. The use of the appropriate function depends on whether the design includes one or more entries. The function for more than one entry is the HSD(). A user enters in the arguments of the HSD() function the number of entries being evaluated (E), the K parameter of the design (K), the number of rows (rows), the number of plants per row (plpr), the plant-to-plant distance in meters (distance) and the display of the HSD in polygons or points (poly = TRUE or FALSE). The first four arguments of the function must enter a value, while the last two (poly and control) are optional and default values are TRUE and FALSE, respectively. If the user enters a wrong number of entries or value of parameter K into the function or omits a required argument, the execution of the command leads to a warning message and appropriate instructions.

To construct the HSD(7) comprising 12 rows of 12 plants (total 144 plants), with planting distances of 1 m and parameter K = 2, the user executes the following command:

```
> HSD(E = 7, K = 2, rows = 12, plpr = 12, d = 1, poly = TRUE, control = FALSE)
or without the names of the arguments.
> HSD(7, 2, 12, 12, 1)
```

The function returns the layout of the plants in the experimental field (Figure 6a,b). It also returns a data frame with six columns (Figure 7). The first column of the data frame represents the entries, the second column represents the rows of the design, the third

column represents the position of the plant within the row, the fourth and fifth columns represent the planting distances, and the last column represents the yield data (at the moment empty). However, in designs with many plants, only a limited number of rows of the data frame are displayed in the command console. To display all the rows of the data frame, the user needs to increase the limit of max.print by changing the global options in R with the following command:

```
> options(max.print = 10,000)
```

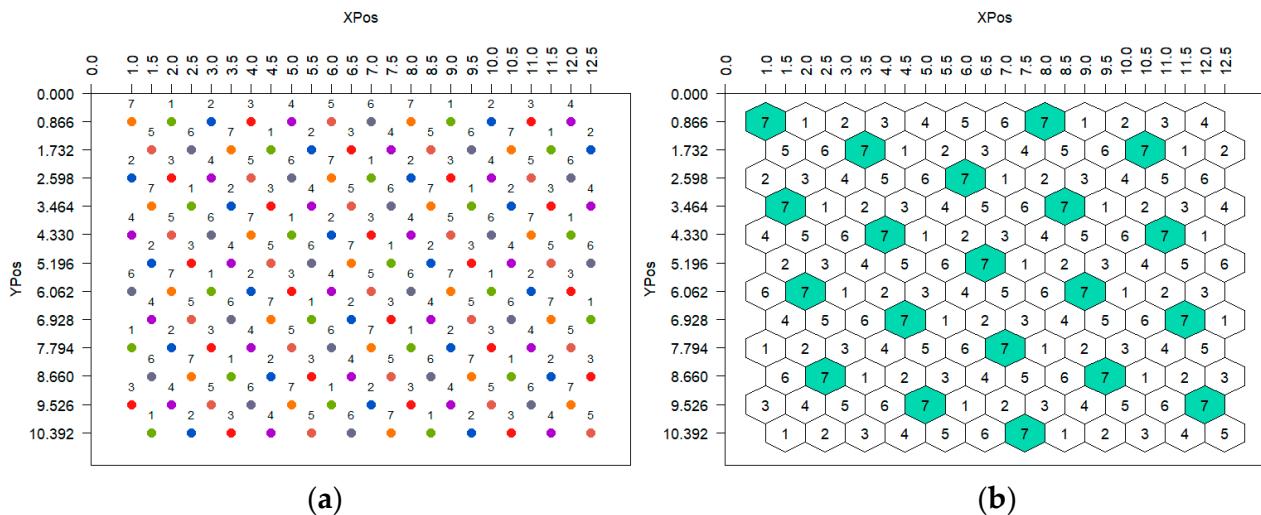


Figure 6. The display of the HSD(7) with points (a) and polygons (b).

Console Terminal Background Jobs

R 4.3.0 ~

```
> HSD(E = 7, K = 2, rows = 12, plpr = 12, distance = 1, poly = TRUE, control = FALSE)
```

Entry	Row	Plant	XPos	YPos	Data
1	7	1	1.0	0.8660254	NA
2	1	2	2.0	0.8660254	NA
3	2	3	3.0	0.8660254	NA
4	3	4	4.0	0.8660254	NA
5	4	5	5.0	0.8660254	NA
6	5	6	6.0	0.8660254	NA
7	6	7	7.0	0.8660254	NA
8	7	1	8.0	0.8660254	NA
9	1	1	9.0	0.8660254	NA
10	2	1	10.0	0.8660254	NA
11	3	1	11.0	0.8660254	NA
12	4	1	12.0	0.8660254	NA
13	5	2	1.5	1.7320508	NA
14	6	2	2.5	1.7320508	NA
15	7	2	3.5	1.7320508	NA
16	1	2	4.5	1.7320508	NA
17	2	2	5.5	1.7320508	NA
18	3	2	6.5	1.7320508	NA
19	4	2	7.5	1.7320508	NA
20	5	2	8.5	1.7320508	NA
21	6	2	9.5	1.7320508	NA

Figure 7. The data frame with plants position for the HSD(7) (21 out of 144 rows).

When the HSD evaluates a single entry, it could be combined with one or three additional entries used as controls. The function for a single-entry HSD without control, called HSD0(), does not require the input of the parameter K. The user enters in the HSD0() function the number of rows (rows), the number of plants per row (plpr), the plant-to-plant distance in meters (d), and the display of the design (poly = TRUE or FALSE). To construct a single-entry HSD comprising 12 rows of 12 plants (total 144 plants), with planting distances of 1 m, the user executes the following command:

```
> HSD0(rows = 12, plpr = 12, d = 1, poly = TRUE)
```

The output of function HSD() includes both a data frame and a plot depicting the layout of plants in the experimental field.

The other two functions corresponding to designs with one or three entry-controls are HSD01() and HSD03() and require the K parameter. To construct a single-entry HSD with one entry-control comprising 12 rows of 12 plants, with planting distance of 1 m and parameter $K = 1$, the user executes the following command:

```
> HSD01(K = 1, rows = 12, plpr = 12, d = 1, poly = TRUE)
```

The HSD01() function returns a data frame and a plot with the layout of the plants in the experimental field (Figure 8).

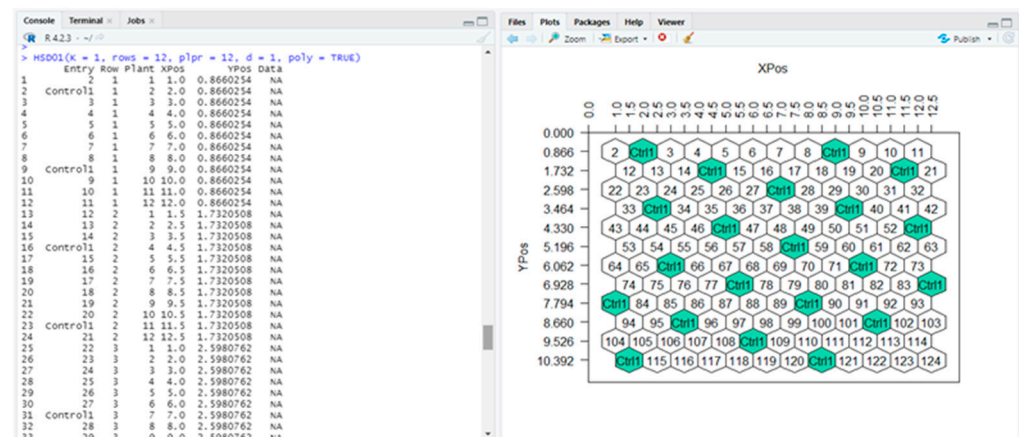


Figure 8. The data frame and the plot of a single-entry HSD with an entry-control ($K = 1$, $X = 2$, $Y = 1$ and $d = 1$ m) (33 out of 144 rows).

To construct a single-entry HSD with three entry-controls comprising 12 rows of 12 plants, with planting distances of 1 m and parameter $K = 1$, the user executes the following command:

```
> HSD03(K = 1, rows = 12, plpr = 12, d = 1, poly = TRUE)
```

The HSD03() function returns a data frame and a plot with the layout of the plants in the experimental field (Figure 9).

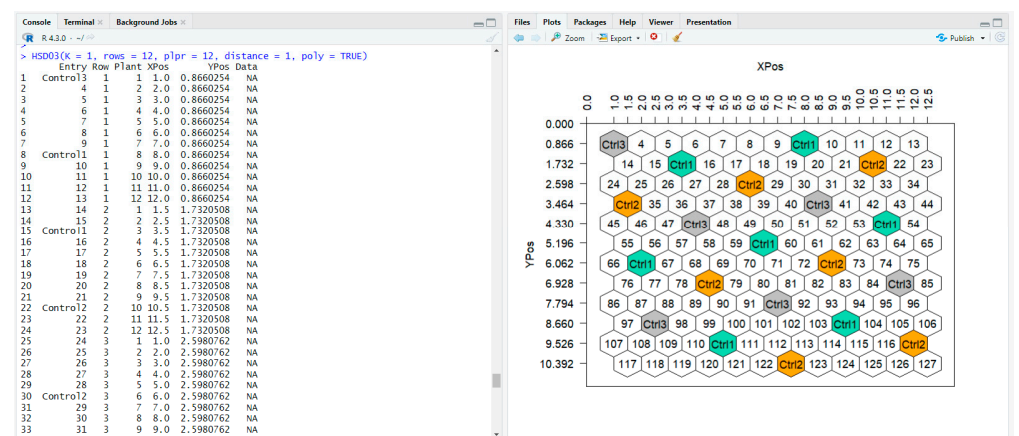


Figure 9. The data frame and the layout of a single-entry HSD with three entries-controls ($K = 1$, $X = 2$, $Y = 1$ and $d = 1$ m) (33 out of 144 rows).

Suppose plant breeders intend to evaluate the plants of a single entry with the addition of an entry-control, whose plants are more sparsely spaced than the standard control density; they should use the HSD() function with the control argument. In this case, the plant breeder will enter in the function HSD() the parameters of an HSD (recommended to be ungrouped) and, in the control argument, will declare TRUE. The result is a multiple-entry honeycomb arrangement, where all the positions correspond to plants of the single

entry except those corresponding to the entry-control (that of the largest code number). The control plants are spaced according to the X and Y coordinates of the chosen HSD(E). For example, the control plants will occupy an area of the experimental field of approximately 7.5% if the HSD(13) is applied ($X = 3$ and $Y = 1$) (Figure 10a) and 5.5% in the case of the HSD(19) ($X = 4$ and $Y = 1$) (Figure 10b).

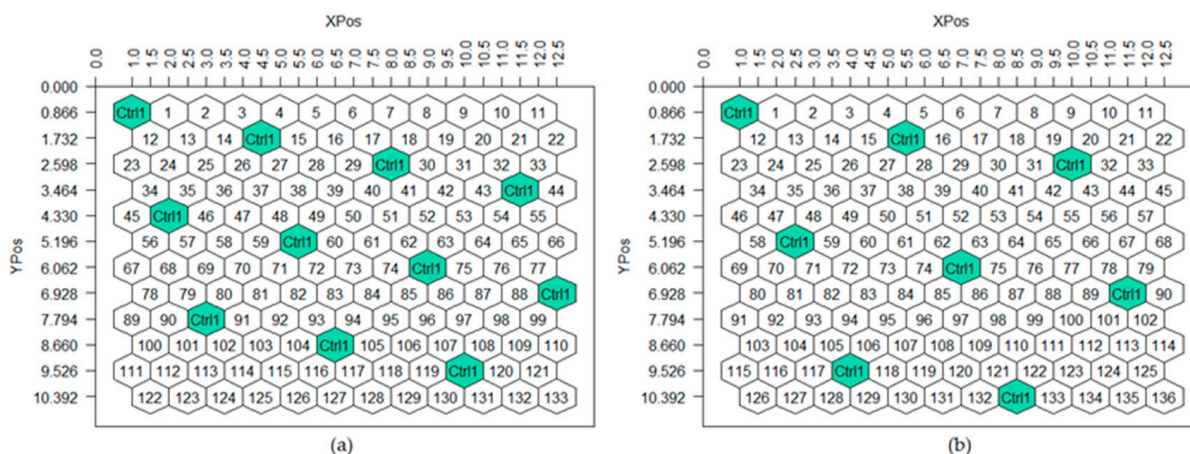


Figure 10. The plots of two single-entry HSDs with control coordinates (a) $X = 3$, $Y = 1$ and (b) $X = 4$, $Y = 1$.

3.4. Import Experimental Data to Data Frame

The execution of the commands for constructing the HSD returns a data frame that includes an empty column for data input named Data. To import the data into the data frame, the user should first assign the data frame as an object and then input the data into a vector and assign it to the object. The experimental data are entered into a vector according to the order they appear in the data frame. If values are missing, they are replaced with the symbol NA.

The steps for constructing the HSD(4) with 4 rows and 4 plants per row and importing the data (16 observations, 2 of which are missing values) are as follows:

Assign the data frame as an object named my_design.

```
> my_design <- HSD(4,1,4,4,1)
```

Data entry into the vector.

```
> my_data <- c(5,8,7,NA,6,9,4,5,8,4,2,NA,6,8,4,5)
```

Assign the data vector into the object.

```
> my_design$Data <- my_data
```

To achieve the most efficient handling of the data frame regarding storage and editing, it is highly recommended to export it to spreadsheet files such as Excel or OpenDocuments using the commands of the package writexl [30]. Upon completing the data import into the spreadsheet file, the file can then be reimported into the R environment.

3.5. Analysis of Honeycomb Selection Designs

The analysis of the experimental data of any HSD is conducted by the function analysis() of the package. The function analysis() contains the arguments: Main_Data_Frame, Response_Vector, circle, blocks, row_element, plant_element and CRS. In order to perform the analysis, it is necessary to declare the first two arguments, while the remaining arguments can be optionally declared. In the argument Main_Data_Frame, the user enters the assigned object data frame generated by one of the functions HSD(), HSD0(), HSD01() and HSD03(), and in the argument Response Vector, the vector of the data column.

The function defaults to the six nearest neighboring plants if the user does not set a specific circle size. Alternatively, instead of circles, the user can use the complete circular block, the moving ring that contains all nearest neighboring entries. In this case, the block

argument should be declared as TRUE. To locate the circle or the complete circular block surrounding a specific plant, the user declares in the argument row element the row number of the plant and, in the argument plant element, the position of the plant in the row. In addition, the user can declare the number of selected plants used to calculate the CRS, considering the entry with the highest identification number as the control.

The function analysis() returns two data frames. The first data frame presents for each plant the following: yield, moving ring size, moving ring mean, PYE value, entry mean, entry number of plants, entry standard deviation, entry HI, and PPE value. The second data frame presents for each entry: number of plants, mean yield, Coefficient of Variation (CV%), standard deviation, HI value, EYI value, EPE value, mPYE value, mPPE value, and CRS value. The function also returns the descriptive statistics of the trial (grand mean, standard deviation, and CV%), the size of the moving ring, and the number of selected plants used to calculate the CRS.

4. Illustrative Examples

The following example is a showcase for the construction and analysis of the HSD. The HSD(31) trial included 31 maize lines and was conducted in 2012 on the farm of the Technological Education Institute of Western Macedonia in Florina, Greece [31]. For construction and analysis, the parameter $K = 5$ was used. The experiment consisted of 30 rows and 32 plants per row (960 total plants), while the planting distance was 1.25 m. To construct the design, the user must initially select the design parameters by executing the function generate(31). Then, for the construction of the HSD(31), the command will be executed:

```
> FLORINA <- HSD(E = 31, K = 5, rows = 30, plpr = 32, d = 1.25, poly = TRUE,
control = FALSE)
```

This command will return the plot with the layout of the plants (Figure 11).

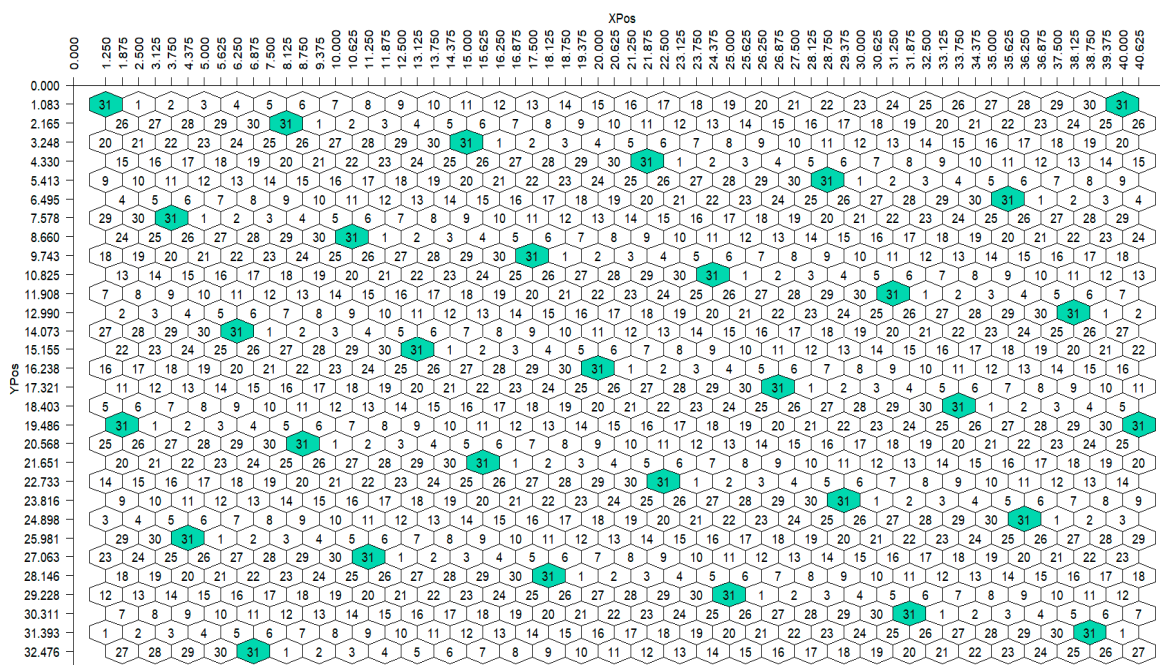
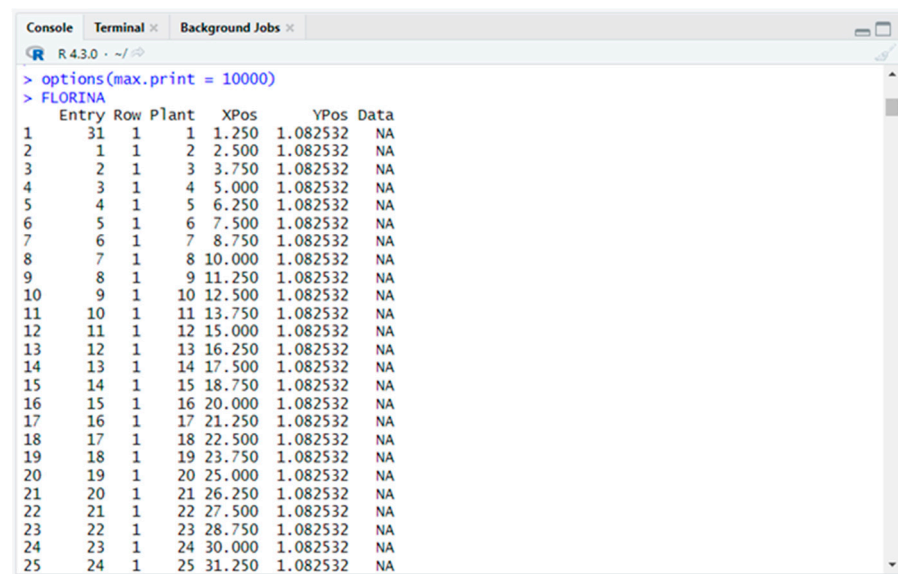


Figure 11. The HSD(31) ($K = 5$, $X = 5$, $Y = 1$ and $d = 1.25$ m).

By executing the object of the above command (FLORINA), the data frame will be displayed (Figure 12).



```

> options(max.print = 10000)
> FLORINA

```

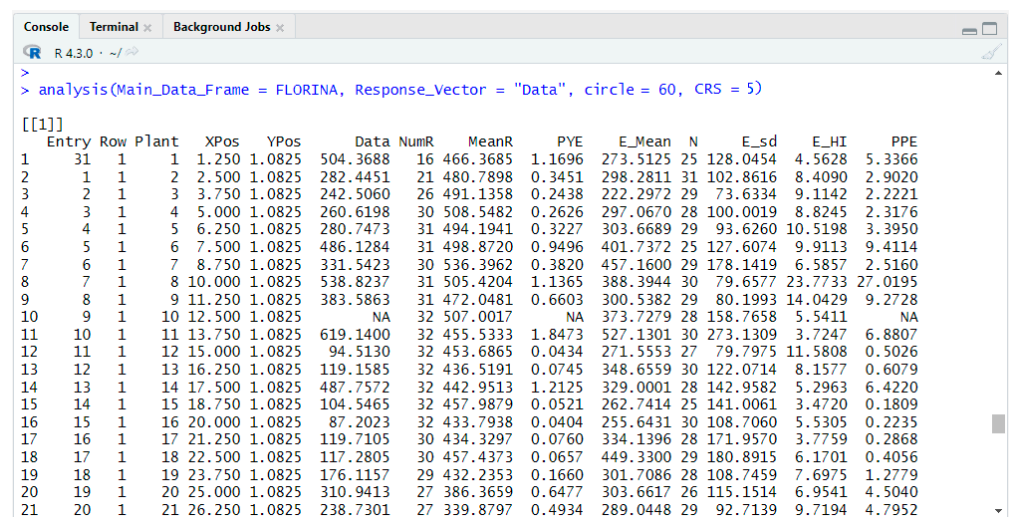
	Entry	Row	Plant	XPos	YPos	Data
1	31	1	1	1.250	1.082532	NA
2	1	1	2	2.500	1.082532	NA
3	2	1	3	3.750	1.082532	NA
4	3	1	4	5.000	1.082532	NA
5	4	1	5	6.250	1.082532	NA
6	5	1	6	7.500	1.082532	NA
7	6	1	7	8.750	1.082532	NA
8	7	1	8	10.000	1.082532	NA
9	8	1	9	11.250	1.082532	NA
10	9	1	10	12.500	1.082532	NA
11	10	1	11	13.750	1.082532	NA
12	11	1	12	15.000	1.082532	NA
13	12	1	13	16.250	1.082532	NA
14	13	1	14	17.500	1.082532	NA
15	14	1	15	18.750	1.082532	NA
16	15	1	16	20.000	1.082532	NA
17	16	1	17	21.250	1.082532	NA
18	17	1	18	22.500	1.082532	NA
19	18	1	19	23.750	1.082532	NA
20	19	1	20	25.000	1.082532	NA
21	20	1	21	26.250	1.082532	NA
22	21	1	22	27.500	1.082532	NA
23	22	1	23	28.750	1.082532	NA
24	23	1	24	30.000	1.082532	NA
25	24	1	25	31.250	1.082532	NA

Figure 12. The data frame of the HSD(31) (25 out of 960 rows).

Data can be added to the data frame either by directly using a data vector or by exporting the dataset to a spreadsheet and then importing the data. The analysis of the data is performed by executing the function `analysis()`, specifying the name of the dataset (e.g., FLORINA), the data column (Data), the circle size (60 plants), as well as the number of selected plants for the CRS (the 5 plants of the highest PYE):

```
> analysis(Main_Data_Frame = FLORINA, Response_Vector = "Data", circle = 60, CRS = 5)
```

The function `analysis()` returns two data frames, one for the evaluation of individual plants (Figure 13) and one for the evaluation of the entries (Figure 14).



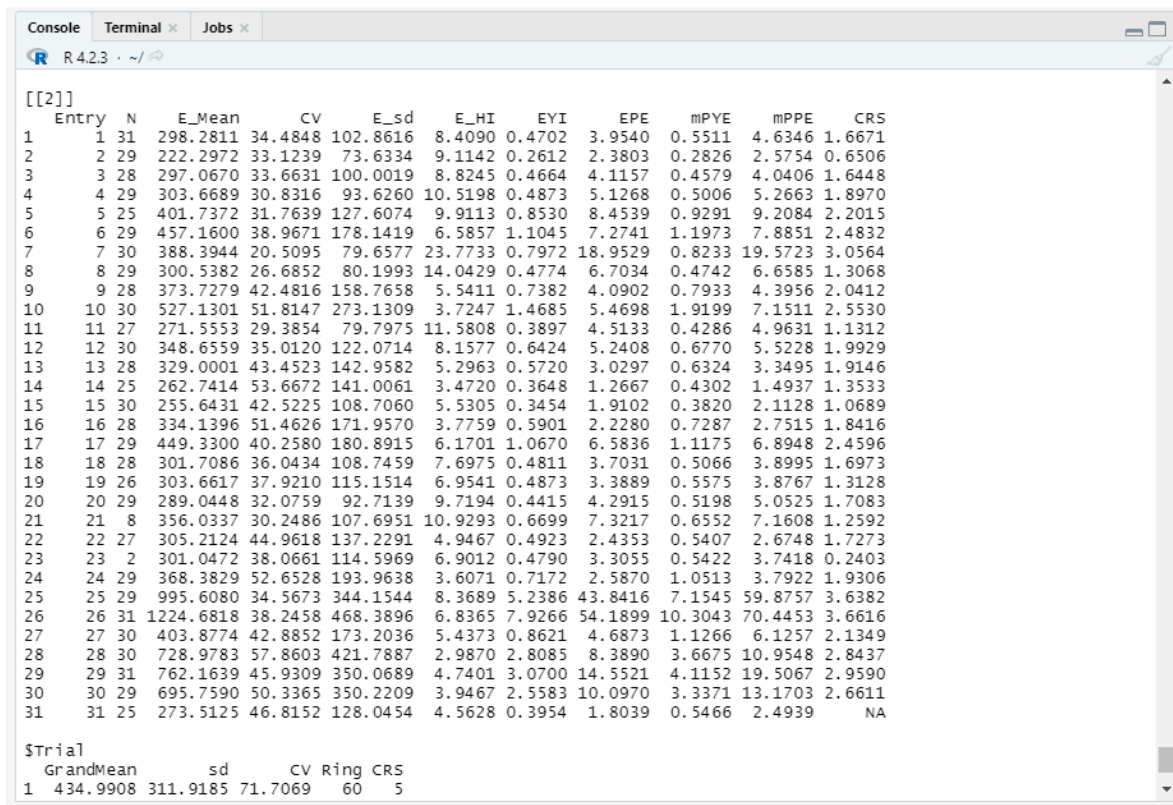
```

> analysis(Main_Data_Frame = FLORINA, Response_Vector = "Data", circle = 60, CRS = 5)

```

	Entry	Row	Plant	XPos	YPos	Data	NumR	MeanR	PYE	E_Mean	N	E_sd	E_HI	PPE
1	31	1	1	1.250	1.0825	504.3688	16	466.3685	1.1696	273.5125	25	128.0454	4.5628	5.3366
2	1	1	2	2.500	1.0825	282.4451	21	480.7898	0.3451	298.2811	31	102.8616	8.4090	2.9020
3	2	1	3	3.750	1.0825	242.5060	26	491.1358	0.2438	222.2972	29	73.6334	9.1142	2.2221
4	3	1	4	5.000	1.0825	260.6198	30	508.5482	0.2626	297.0670	28	100.0019	8.8245	2.3176
5	4	1	5	6.250	1.0825	280.7473	31	494.1941	0.3227	303.6689	29	93.6260	10.5198	3.3950
6	5	1	6	7.500	1.0825	486.1284	31	498.8720	0.9496	401.7372	25	127.6074	9.9113	9.4114
7	6	1	7	8.750	1.0825	331.5423	30	536.3962	0.3820	457.1600	29	178.1419	6.5857	2.5160
8	7	1	8	10.000	1.0825	538.8237	31	505.4204	1.1365	388.3944	30	79.6577	23.7733	27.0195
9	8	1	9	11.250	1.0825	383.5863	31	472.0481	0.6603	300.5382	29	80.1993	14.0429	9.2728
10	9	1	10	12.500	1.0825	NA	32	507.0017	NA	373.7279	28	158.7658	5.5411	NA
11	10	1	11	13.750	1.0825	619.1400	32	455.5333	1.8473	527.1301	30	273.1309	3.7247	6.8807
12	11	1	12	15.000	1.0825	94.5130	32	453.6865	0.0434	271.5553	27	79.7975	11.5808	0.5026
13	12	1	13	16.250	1.0825	119.1585	32	436.5191	0.0745	348.6559	30	122.0714	8.1577	0.6079
14	13	1	14	17.500	1.0825	487.7572	32	442.9513	1.2125	329.0001	28	142.9582	5.2963	6.4220
15	14	1	15	18.750	1.0825	104.5465	32	457.9879	0.0521	262.7414	25	141.0061	3.4720	0.1809
16	15	1	16	20.000	1.0825	87.2023	32	433.7938	0.0404	255.6431	30	108.7060	5.5305	0.2235
17	16	1	17	21.250	1.0825	119.7105	30	434.3297	0.0760	334.1396	28	171.9570	3.7759	0.2868
18	17	1	18	22.500	1.0825	117.2805	30	457.4373	0.0657	449.3300	29	180.8915	6.1701	0.4056
19	18	1	19	23.750	1.0825	176.1157	29	432.2353	0.1660	301.7086	28	108.7459	7.6975	1.2779
20	19	1	20	25.000	1.0825	310.9413	27	386.3659	0.6477	303.6617	26	115.1514	6.9541	4.5040
21	20	1	21	26.250	1.0825	238.7301	27	339.8797	0.4934	289.0448	29	92.7139	9.7194	4.7952

Figure 13. The evaluation of the individual plants (21 out of 960 rows).



```
[[2]]
```

	Entry	N	E_Mean	CV	E_sd	E_HI	EYI	EPE	mPYE	mPPE	CRS
1	1	31	298.2811	34.4848	102.8616	8.4090	0.4702	3.9540	0.5511	4.6346	1.6671
2	2	29	222.2972	33.1239	73.6334	9.1142	0.2612	2.3803	0.2826	2.5754	0.6506
3	3	28	297.0670	33.6631	100.0019	8.8245	0.4664	4.1157	0.4579	4.0406	1.6448
4	4	29	303.6689	30.8316	93.6260	10.5198	0.4873	5.1268	0.5006	5.2663	1.8970
5	5	25	401.7372	31.7639	127.6074	9.9113	0.8530	8.4539	0.9291	9.2084	2.2015
6	6	29	457.1600	38.9671	178.1419	6.5857	1.1045	7.2741	1.1973	7.8851	2.4832
7	7	30	388.3944	20.5095	79.6577	23.7733	0.7972	18.9529	0.8233	19.5723	3.0564
8	8	29	300.5382	26.6852	80.1993	14.0429	0.4774	6.7034	0.4742	6.6585	1.3068
9	9	28	373.7279	42.4816	158.7658	5.5411	0.7382	4.0902	0.7933	4.3956	2.0412
10	10	30	527.1301	51.8147	273.1309	3.7247	1.4685	5.4698	1.9199	7.1511	2.5530
11	11	27	271.5553	29.3854	79.7975	11.5808	0.3897	4.5133	0.4286	4.9631	1.1312
12	12	30	348.6559	35.0120	122.0714	8.1577	0.6424	5.2408	0.6770	5.5228	1.9929
13	13	28	329.0001	43.4523	142.9582	5.2963	0.5720	3.0297	0.6324	3.3495	1.9146
14	14	25	262.7414	53.6672	141.0061	3.4720	0.3648	1.2667	0.4302	1.4937	1.3533
15	15	30	255.6431	42.5225	108.7060	5.5305	0.3454	1.9102	0.3820	2.1128	1.0689
16	16	28	334.1396	51.4626	171.9570	3.7759	0.5901	2.2280	0.7287	2.7515	1.8416
17	17	29	449.3300	40.2580	180.8915	6.1701	1.0670	6.5836	1.1175	6.8948	2.4596
18	18	28	301.7086	36.0434	108.7459	7.6975	0.4811	3.7031	0.5066	3.8995	1.6973
19	19	26	303.6617	37.9210	115.1514	6.9541	0.4873	3.3889	0.5575	3.8767	1.3128
20	20	29	289.0448	32.0759	92.7139	9.7194	0.4415	4.2915	0.5198	5.0525	1.7083
21	21	8	356.0337	30.2486	107.6951	10.9293	0.6699	7.3217	0.6552	7.1608	1.2592
22	22	27	305.2124	44.9618	137.2291	4.9467	0.4923	2.4353	0.5407	2.6748	1.7273
23	23	2	301.0472	38.0661	114.5969	6.9012	0.4790	3.3055	0.5422	3.7418	0.2403
24	24	29	368.3829	52.6528	193.9638	3.6071	0.7172	2.5870	1.0513	3.7922	1.9306
25	25	29	995.6080	34.5673	344.1544	8.3689	5.2386	43.8416	7.1545	59.8757	3.6382
26	26	31	1224.6818	38.2458	468.3896	6.8365	7.9266	54.1899	10.3043	70.4453	3.6616
27	27	30	403.8774	42.8852	173.2036	5.4373	0.8621	4.6873	1.1266	6.1257	2.1349
28	28	30	728.9783	57.8603	421.7887	2.9870	2.8085	8.3890	3.6675	10.9548	2.8437
29	29	31	762.1639	45.9309	350.0689	4.7401	3.0700	14.5521	4.1152	19.5067	2.9590
30	30	29	695.7590	50.3365	350.2209	3.9467	2.5583	10.0970	3.3371	13.1703	2.6611
31	31	25	273.5125	46.8152	128.0454	4.5628	0.3954	1.8039	0.5466	2.4939	NA

```
$Trial
```

	GrandMean	sd	CV	Ring	CRS
1	434.9908	311.9185	71.7069	60	5

Figure 14. The evaluation of the 31 entries.

From the first data frame, the plant breeder ranks the plants in descending order based on the plant yield efficient index (PYE).

From the second data frame, the plant breeder ranks the entries in descending order based on one of the following measures, depending on their nature (highly heterogeneous or fairly homogeneous): yield performance (mean or EYI or mPYE), stability (HI), combined yield and stability (EPE or mPPE). Single-plant selection for high PYE (the first data frame) applies within the top entries. If experimentation focuses on evaluating germplasms as potential starting material for breeding, entries of the second frame are ranked on the Coefficient of Response to Selection (CRS). In the present example, CRS values were calculated using entry 31 as the control.

5. Conclusions

This paper gives a detailed presentation of the upgraded version of the ‘rhoneycomb’, a free and open-source R package comprising functions that can enhance the construction, visualization, and analysis of HSDs for plant breeders, providing them with a more precise and user-friendly approach. The upgraded package offers additional functions and input options to customize further their research aims. The functionalities provided by the ‘rhoneycomb’ package enable the selection of high-yielding genotypes and ensure stability in trait performance. These functionalities increase efficiency and speed, aid the mechanization of the early stages of plant selection, and serve as a valuable contribution to the crop breeding community, offering efficient assistance in the initial stages of a breeding program. Thus, the proposed new version of the package can greatly impact crop breeding programs by making the selection of superior plants more efficient and reliable. The package is free to download at CRAN: <https://cran.r-project.org/web/packages/rhoneycomb/> (accessed on 26 June 2023) with additional examples and information.

Author Contributions: Conceptualization, A.K. and P.J.B.; methodology, A.K., I.T. and P.J.B.; software, A.K., N.A. and M.-G.G.; validation, A.K., E.T., I.T. and P.J.B.; formal analysis, A.K. and I.T.; investigation, A.K., N.A. and M.-G.G.; resources, P.J.B.; data curation, A.K. and I.T.; writing—original draft preparation, A.K., N.A. and M.-G.G.; writing—review and editing, E.T., P.J.B. and I.T.; visualization, A.K. and N.A.; supervision, I.T. and P.J.B.; project administration, P.J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ahmar, S.; Gill, R.A.; Jung, K.-H.; Faheem, A.; Qasim, M.U.; Mubeen, M.; Zhou, W. Conventional and Molecular Techniques from Simple Breeding to Speed Breeding in Crop Plants: Recent Advances and Future Outlook. *Int. J. Mol. Sci.* **2020**, *21*, 2590. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zystro, J.; Colley, M.; Dawson, J. Alternative Experimental Designs for Plant Breeding. In *Plant Breeding Reviews*; Goldman, I., Ed.; Wiley: Hoboken, NJ, USA, 2018. [\[CrossRef\]](#)
- Cullis, B.R.; Smith, A.B.; Cocks, N.A. The Design of Early-Stage Plant Breeding Trials Using Genetic Relatedness. *JABES* **2020**, *25*, 553–578. [\[CrossRef\]](#)
- Gonçalves, E.; St Aubyn, A.; Martins, A. Experimental designs for evaluation of genetic variability and selection of ancient grapevine varieties: A simulation study. *Heredity* **2010**, *104*, 552–562. [\[CrossRef\]](#) [\[PubMed\]](#)
- Federer, W.T.; Crossa, J.I. Screening Experimental Designs for Quantitative Trait Loci, Association Mapping, Genotype-by-Environment Interaction, and Other Investigations. *Front. Physiol.* **2012**, *3*, 156. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mathews, K.L.; Crossa, J. *Experimental Design for Plant Improvement*; Springer: Berlin/Heidelberg, Germany, 2022. [\[CrossRef\]](#)
- Fasoulas, A.C. *A New Approach to Breeding Superior Yielding Varieties*; Dept. Genetics and Plant Breeding, Aristotelian University of Thessaloniki: Thessaloniki, Greece, 1973; Volume 3, p. 42.
- Fasoulas, A.C. *Principles and Methods of Plant Breeding*; Dept. Gen. Plant Breeding, Aristotelian University of Thessaloniki: Thessaloniki, Greece, 1981; Volume 11, p. 147.
- Fasoulas, A.C. *The Honeycomb Methodology of Plant Breeding*; Aristotelian University of Thessaloniki: Thessaloniki, Greece, 1988.
- Fasoula, D.A. Correlations between auto-, allo- and nil-competition and their implications in plant breeding. *Euphytica* **1990**, *50*, 57–62. [\[CrossRef\]](#)
- Falconer, D.S. *Introduction to Quantitative Genetics*, 3rd ed.; Longman Scientific and Terminal Wiley: New York, NY, USA, 1989.
- Uphoff, N.; Fasoula, V.A.; Iswandi, A.; Kassam, A.; Thakur, A.K. Improving the phenotypic expression of rice genotypes: Rethinking intensification for production systems and selection practices for rice breeding. *Crop J.* **2015**, *3*, 174–189. [\[CrossRef\]](#)
- Fasoula, V.A.; Tokatlidis, I.S. Development of crop cultivars by honeycomb breeding. *Agron. Sustain. Dev.* **2012**, *32*, 161–180. [\[CrossRef\]](#)
- Tokatlidis, I. Crop resilience via inter-plant spacing brings to the fore the productive ideotype. *Front. Plant Sci.* **2022**, *13*, 934359. [\[CrossRef\]](#) [\[PubMed\]](#)
- Tokatlidis, I. Sampling the spatial heterogeneity of the honeycomb model in maize and wheat breeding trials: Analysis of secondary data compared to popular classical designs. *Exp. Agric.* **2016**, *52*, 371–390. [\[CrossRef\]](#)
- Fasoula, V.A.; Fasoula, D.A. Principles underlying genetic improvement for high and stable crop yield potential. *Field Crops Res.* **2002**, *75*, 191–209. [\[CrossRef\]](#)
- Fasoula, V.A.; Fasoula, D.A. Honeycomb breeding: Principles and applications. In *Plant Breeding Reviews*; Janick, J., Ed.; John Wiley and Sons Inc.: New York, NY, USA, 2000; Volume 18, pp. 177–250. [\[CrossRef\]](#)
- Fischer, R.A. Breeding wheat for increased potential yield: Contrasting ideas from donald and fasoulas, and the case for early generation selection under nil competition. *Field Crops Res.* **2020**, *252*, 107782. [\[CrossRef\]](#)
- Fasoula, V.A.; Thompson, K.C.; Mauromoustakos, A. The prognostic breeding application JMP add-in program. *Agronomy* **2019**, *9*, 25. [\[CrossRef\]](#)
- Katsileros, A.; Antonetsis, N.; Gkika, M.G.; Tani, E.; Tokatlidis, I.; Bebeli, P.J. rhoneycomb: An R package for the construction and analysis of honeycomb selection designs. *Softw. Impacts* **2023**, *16*, 100490. [\[CrossRef\]](#)
- Fasoulas, A.C.; Fasoula, V.A. Honeycomb selection designs. In *Plant Breeding Reviews*; Janick, J., Ed.; John Wiley and Sons Inc.: New York, NY, USA, 1995; Volume 13, pp. 87–139. [\[CrossRef\]](#)
- Katsileros, A.; Antonetsis, N.; Gkika, M.G.; Tokatlidis, I.; Bebeli, P.J. Evaluation of Greek varieties of dread wheat in terms of yield and its components, with two experimental arrangements. In Proceedings of the 18th Conference of the Hellenic Scientific Society of Genetics and Plant Breeding, Volos, Greece, 5–7 October 2022.

23. Fasoula, V.A. Prognostic breeding: A new paradigm for crop improvement. In *Plant Breeding Reviews*; Janick, J., Ed.; John Wiley and Sons Inc.: New York, NY, USA, 2013; Volume 37, pp. 297–347.
24. Tokatlidis, I.S. Crop adaptation to density to optimise grain yield: Breeding implications. *Euphytica* **2017**, *213*, 92. [CrossRef]
25. Ninou, E.G.; Mylonas, I.G.; Tsivelikas, A.; Ralli, P.; Dordas, C.; Tokatlidis, I.S. Wheat landraces are better qualified as potential gene pools at ultraspaced rather than densely grown conditions. *Sci. World J.* **2014**, *2014*, 957472. [CrossRef]
26. Tokatlidis, I.; Vlachostergios, D. Sustainable Stewardship of the Landrace Diversity. *Diversity* **2016**, *8*, 29. [CrossRef]
27. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: <https://www.Rproject.org/> (accessed on 15 March 2023).
28. Wickham, H.; Hester, J.; Chang, W.; Bryan, J. Devtools: Tools to Make Developing R Packages Easier. 2022. Available online: <https://devtools.r-lib.org/> (accessed on 11 October 2022).
29. RStudio an Integrated Development Environment for, R. RStudio Desktop. 2022. Available online: <https://www.rstudio.com/products/rstudio/download> (accessed on 19 April 2022).
30. Ooms, J.; McNamara, J. writexl: Export Data Frames to Excel 'xlsx' Format, 1.4.2. Available online: <https://cran.r-project.org/web/packages/writexl> (accessed on 6 January 2023).
31. Tokatlidis, I.S.; Dordas, C.; Papathanasiou, F.; Papadopoulos, I.; Pankou, C.; Gekas, F.; Ninou, E.; Mylonas, I.; Tzantarmas, C.; Petrevska, J.-K.; et al. Improved plant yield efficiency is essential for maize rainfed production. *Agron. J.* **2015**, *107*, 1011–1018. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.